(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 December 2022 (22.12.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/266511 A8

(51) International Patent Classification:
C12N 15/62 (2006.01) *A61K 39/215* (2006.01)
C12N 15/85 (2006.01) *A61K 39/295* (2006.01)
C12N 7/00 (2006.01) *C07K 14/005* (2006.01)
A61K 39/00 (2006.01) *A61K 31/7088* (2006.01)
A61K 39/12 (2006.01) *A61P 31/14* (2006.01)
A61K 39/145 (2006.01)

(21) International Application Number:
PCT/US2022/034104

(22) International Filing Date:
17 June 2022 (17.06.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/211,974 17 June 2021 (17.06.2021) US
63/240,278 02 September 2021 (02.09.2021) US
63/275,398 03 November 2021 (03.11.2021) US

(71) Applicant: **ELIXIRGEN THERAPEUTICS, INC.** [US/US]; 855 North Wolfe Street, Suite 624, Baltimore, Maryland 21205 (US).

(72) Inventor: **KO, Minoru S., H.**; c/o Elixirgen Therapeutics, Inc., 855 North Wolfe Street, Suite 624, Baltimore, Maryland 21205 (US).

(74) Agent: **LEKUTIS, Christine** et al.; Morrison & Foerster LLP, 425 Market Street, San Francisco, California 94105 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,

(54) Title: TEMPERATURE-CONTROLLABLE, SELF-REPLICATING RNA VACCINES FOR VIRAL DISEASES

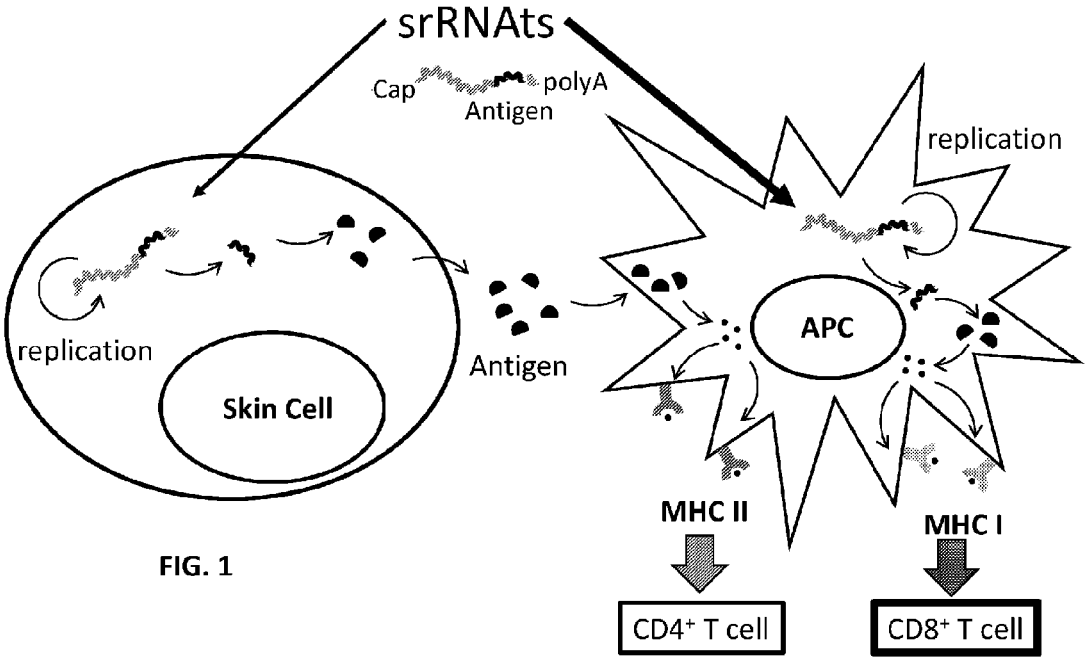


FIG. 1

(57) Abstract: The present disclosure relates to mRNA, self-replicating RNA, and temperature-sensitive, self-replicating RNA encoding a coronavirus nucleocapsid protein or an influenza virus nucleocapsid protein in operable combination with a mammalian signal peptide. The present disclosure relates to mRNA, self-replicating RNA, and temperature-sensitive, self-replicating RNA encoding other viral nucleocapsid protein(s) in operable combination with a mammalian signal peptide. The RNA constructs are suitable for active immunization against a virus in a mammalian subject, such as a human subject.

WO 2022/266511 A8

[Continued on next page]

KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*
— *with information concerning authorization of rectification of an obvious mistake under Rule 91.3 (b) (Rule 48.2(i))*

(48) Date of publication of this corrected version:
09 November 2023 (09.11.2023)

(15) Information about Correction:
see Notice of 09 November 2023 (09.11.2023)

Previous Correction:
see Notice of 13 July 2023 (13.07.2023)

TEMPERATURE-CONTROLLABLE, SELF-REPLICATING RNA VACCINES FOR VIRAL DISEASES

CROSS-REFERENCE TO RELATED APPLICATION(S)

[0001] This application claims the benefit of U.S. Provisional Application No. 63/275,398, filed November 3, 2021, U.S. Provisional Application No. 63/240,278, filed September 2, 2021, and U.S. Provisional Application No. 63/211,974, filed June 17, 2021, each of which is hereby incorporated by reference in its entirety.

SUBMISSION OF SEQUENCE LISTING ON ASCII TEXT FILE

[0002] The content of the following submission on ASCII text file is incorporated herein by reference in its entirety: a computer readable form (CRF) of the Sequence Listing (file name: 699442001440SEQLIST.TXT, date recorded: June 16, 2022, size: 126,113 bytes).

FIELD

[0003] The present disclosure relates to mRNA, self-replicating RNA, and temperature-sensitive, self-replicating RNA encoding a coronavirus nucleocapsid protein or an influenza virus nucleocapsid protein in operable combination with a mammalian signal peptide. The present disclosure relates to mRNA, self-replicating RNA, and temperature-sensitive, self-replicating RNA encoding other viral nucleocapsid protein(s) in operable combination with a mammalian signal peptide. The RNA constructs are suitable for active immunization against a virus in a mammalian subject, such as a human subject.

BACKGROUND

[0004] The betacoronavirus genus encompasses Severe Acute Respiratory Syndrome (SARS)-CoV-2, which caused the COVID-19 pandemic, SARS-CoV-1, which caused the 2002–2004 SARS outbreak, and Middle East Respiratory Syndrome (MERS)-CoV. The COVID-19 pandemic has made design and production of vaccines an urgent necessity for immunization of a large global population.

[0005] The SARS-CoV-2 vaccines currently approved by the U.S. Food & Drug Administration are designed to elicit neutralizing antibodies (nAb) against the Spike (S) protein

or the receptor binding domain (RBD) of the S protein in advance of infection. However, this approach poses a great challenge in that the S protein is not well conserved even between SARS-CoV-1 and SARS-CoV-2 strains. In particular, small amino acid changes that occur among variants often result in conformational changes to the S protein that may significantly reduce the effectiveness of nAb elicited by the specific S protein of the COVID-19 vaccine.

[0006] Continued vaccine development targeting only the betacoronavirus S protein is therefore contemplated to follow the path of seasonal influenza vaccines. This means that the continual emergence of variants will likely require development and production of new vaccines on a periodic basis. Although annual production of betacoronavirus vaccines may be technically feasible, global vaccination efforts involving annual administration of new vaccines are economically and logistically impractical. The problems posted by annual administration of new vaccines present especially undue burdens for low- and middle-income countries.

[0007] Accordingly, there is a need in the art for betacoronavirus vaccines that safely induce a long-lived, immune response that is broadly reactive against SARS-CoV-2 variants. Preferably the long-lived, immune response is broadly reactive with other betacoronaviruses, which cause disease in humans. There is also a need in the art for influenza virus vaccines that are safe and effective in inducing a broadly reactive immune response against influenza A and/or influenza B viruses.

BRIEF SUMMARY

[0008] The present disclosure relates to the use of nucleoproteins (also referred to herein as nucleocapsid proteins) from betacoronaviruses as a vaccine antigen to induce cellular immune responses that are broadly reactive with betacoronavirus variants. In some embodiments, a temperature-controllable, self-replicating RNA (referred to herein as srRNAts and c-srRNA) vaccine platform is utilized. The c-srRNA vaccine platform is advantageous for induction of a potent cellular immune response after intradermal administration. In some embodiments, a nucleoprotein from SARS-CoV-2 is expressed in host cells to address infection by both SARS-CoV-2 and SARS-CoV-1, as well as variants thereof. In some embodiments, a nucleoprotein from a coronavirus is fused with a signal peptide of the human CD5 antigen and expressed in host cells to enhance the cellular immune response elicited against the coronavirus. In some

embodiments, a nucleoprotein from a first coronavirus is fused to a nucleoprotein from a second coronavirus, which is different from the first coronavirus. In some embodiments, the fusion protein comprises a tandem array of two or three coronavirus nucleoproteins. In a subset of these embodiments, the fusion protein comprises a SARS-CoV-2 nucleoprotein and a MERS-CoV nucleoprotein. In some embodiments, the fusion protein further comprises a coronavirus spike protein or fragment thereof. In this way, a more broadly reactive coronavirus-specific immune response is stimulated.

[0009] The present disclosure also relates to the use of nucleoproteins (also referred to herein as nucleocapsid proteins) from influenza viruses as a vaccine antigen to induce cellular immune responses that are broadly reactive with influenza A and/or influenza B viruses, which rapidly change over time as a consequence of antigen drift and antigen shift. In some embodiments, a temperature-controllable, self-replicating RNA vaccine platform is utilized. The c-srRNA vaccine platform is advantageous for induction of a potent cellular immune response after intradermal administration. In some embodiments, a nucleoprotein from one subtype of influenza A (FluA) virus is expressed in host cells to address infection by the same and different subtypes of FluA. In some embodiments, a nucleoprotein from one lineage of influenza B (FluB) virus is expressed in host cells to address infection by the same and different lineages of FluB. In some embodiments, a nucleoprotein from an influenza virus is fused with a signal peptide of the human CD5 antigen and expressed in host cells to enhance the cellular immune response elicited against the influenza virus. In some embodiments, a nucleoprotein from a FluA virus is fused to a nucleoprotein from a FluB virus. In some embodiments, the fusion protein comprises a tandem array of two or three nucleoproteins from one or more strains of FluA and/or one or more lineages of FluB. In some embodiments, the fusion protein further comprises an influenza hemagglutinin or fragment thereof. In this way, a more broadly reactive influenza-specific immune response is stimulated.

[0010] The present disclosure also relates to the use of nucleoproteins (also referred to herein as nucleocapsid proteins) from ebolaviruses as a vaccine antigen to induce cellular immune responses that are broadly reactive with two, three or four species of ebolavirus that infect humans. In some embodiments, a temperature-controllable, self-replicating RNA vaccine platform is utilized. The c-srRNA vaccine platform is advantageous for induction of a potent

cellular immune response after intradermal administration. In some embodiments, a nucleoprotein from an ebolavirus is fused with a signal peptide of the human CD5 antigen and expressed in host cells to enhance the cellular immune response elicited against the ebolavirus. In some embodiments, a nucleoprotein from a first ebolavirus species is fused to a nucleoprotein from a second ebolavirus species, which is optionally fused to a nucleoprotein of a third ebolavirus species, which is optionally fused to a nucleoprotein of a fourth ebolavirus species. In some embodiments, the fusion protein comprises a tandem array of two, three or four nucleoproteins or fragments thereof from two or more species of ebolavirus. In some embodiments, the fusion protein further comprises an ebolavirus envelope glycoprotein or fragment thereof. In this way, a more broadly reactive ebolavirus-specific immune response is stimulated.

[0011] Among other embodiments, the present disclosure provides compositions comprising an excipient and a temperature-controllable, self-replicating RNA. In some embodiments, the composition comprises a chitosan. In some embodiments, the chitosan is a low molecular weight (about 3-5 kDa) chitosan oligosaccharide, such as chitosan oligosaccharide lactate. In some embodiments, the composition does not comprise liposomes or lipid nanoparticles.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] **FIG. 1** shows a schematic of the mechanism for induction of cellular (CD4+ and CD8+ T cell) immune responses after intradermal injection of a temperature-controllable, self-replicating RNA (referred to herein as srRNAts and c-srRNA) vaccine.

[0013] **FIG. 2** shows a schematic diagram of SARS-CoV-2 nucleocapsid (N) proteins expressed from mRNA, self-replicating RNA, or temperature-sensitive, self-replicating RNA (srRNAts) delivered to mammalian host cells. In exemplary embodiments, the coding region of the N protein is the gene of interest (GOI) inserted within the srRNAts. The amino acid sequence of the G5004 antigen is set forth as SEQ ID NO:5. The G5004 antigen is a SARS-CoV-2 N protein devoid of a signal peptide. The amino acid sequence of the G5005 antigen is set forth as SEQ ID NO:6. The G5005 antigen is a fusion protein comprising the signal peptide sequence from the human CD5 antigen (CD5-SP) set forth as SEQ ID NO:8, and a SARS-CoV-2

N protein, in which the CD5-SP replaces the start methionine at position 1 of the N protein. The amino acid sequence of the G5006 antigen is set forth as SEQ ID NO:7. The G5006 antigen is a fusion protein comprising the signal peptide sequence from CD5-SP, a SARS-CoV-2 N protein, and a MERS-CoV N protein. The nucleotide sequence encoding the G5004 antigen is set forth as SEQ ID NO:1. The nucleotide sequence encoding the G5005 antigen is set forth as SEQ ID NO:2. The nucleotide sequence encoding the G5006 antigen is set forth as SEQ ID NO:3, and as a codon-optimized version in SEQ ID NO:4.

[0014] **FIG. 3** shows a schematic diagram of an exemplary method for stimulating an immune response against a coronavirus in a human subject. A temperature-sensitive agent (ts-agent) such as a srRNAts is functional at a permissive temperature, but is non-functional at a non-permissive temperature. The temperature at or just below the surface of a human subject's body (surface body temperature) is a permissive temperature, while the human subject's core body temperature is a higher, non-permissive temperature. Thus, a ts-agent administered intradermally to the human subject is functional while located at the permissive temperature just below the surface of the human subject's body.

[0015] **FIG. 4A** and **FIG. 4B** show the frequency of cytokine-secreting cells in samples of splenocytes obtained from CD-1 outbred mice that had been immunized by a single intradermal injection of 100 μL solution containing either 5 μg or 25 μg of a temperature-controllable self-replicating RNA (srRNA1ts2 [PCT/US20/67506]) encoding the G5004 antigen or a placebo (PBO: buffer only). **FIG 4A** shows the frequency of interferon-gamma (INF-γ) spot-forming cells (SFC) and **FIG. 4B** shows the frequency of interleukin-4 (IL-4) SFC in 1x10^6 splenocytes after restimulation by culturing the splenocytes in the presence or absence of a pool of SARS-CoV-2 nucleoprotein peptides. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background). The average and standard deviation (error bars) of five mice (n=5) are shown for each group. Splenocytes were isolated 14 days after intradermal injection.

[0016] **FIG. 5A** and **FIG. 5B** show the frequency of cytokine-secreting cells in samples of splenocytes obtained from CD-1 outbred mice that had been immunized by a single intradermal injection of 100 μL solution containing either 5 μg or 25 μg of a temperature-controllable self-replicating RNA (srRNA1ts2 [PCT/US20/67506]) encoding the G5005 antigen

or a placebo (PBO: buffer only). **FIG 5A** shows the frequency of interferon-gamma (INF-γ) spot-forming cells (SFC) and **FIG. 5B** shows the frequency of interleukin-4 (IL-4) SFC in 1x10^6 splenocytes after restimulation by culturing the splenocytes in the presence or absence of a pool of SARS-CoV-2 nucleoprotein peptides. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background). The average and standard deviation (error bars) of five mice (n=5) are shown for each group. Splenocytes were isolated 14 days after intradermal injection.

[0017] **FIG. 6A** and **FIG. 6B** show the frequency of cytokine-secreting cells in samples of splenocytes obtained from BALB/c mice that had been immunized by a single intradermal injection of 100 μL solution containing either 5 μg or 25 μg of a temperature-controllable self-replicating RNA (srRNA1ts2 [PCT/US20/67506]) encoding the G5005 antigen or a placebo (PBO: buffer only). **FIG 6A** shows the frequency of interferon-gamma (INF-γ) spot-forming cells (SFC) and **FIG. 6B** shows the frequency of interleukin-4 (IL-4) SFC in 1x10^6 splenocytes after restimulation by culturing the splenocytes in the presence or absence of a pool of SARS-CoV-2 nucleoprotein peptides. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background). The average and standard deviation (error bars) of five mice (n=5) are shown for each group. Splenocytes were isolated 30 days after the vaccination.

[0018] **FIG. 7** show the levels of SARS-CoV-2 antigen-reactive immunoglobulin G (IgG) in serum of BALB/c mice that had been immunized by a single intradermal injection of a 100 μL solution containing either 5 μg or 25 μg of a temperature-controllable self-replicating RNA (srRNA1ts2 [PCT/US20/67506]) encoding the G5005 antigen or a placebo (PBO: buffer only). The IgG levels are represented by OD450 in the ELISA. The IgG levels before (Day -1) and after (Day 30) vaccination (Day 0) are shown. The average and standard deviation (error bars) of five mice (n=5) are shown for each group.

[0019] **FIG. 8** shows the frequency of interferon-gamma (INF-γ)-secreting cells in samples of splenocytes obtained from BALB/c mice that had been immunized by a single intradermal injection of 100 μL solution containing either 5 μg or 25 μg of a temperature-controllable self-replicating RNA (srRNA1ts2 [PCT/US20/67506]) encoding the G5006 antigen or a placebo (PBO: buffer only). Specifically, **FIG. 8** shows the frequency of INF-γ spot-forming

cells (SFC) in 1x10^6 splenocytes after restimulation by culturing the splenocytes in the presence or absence of a pool of SARS-CoV-2 nucleoprotein peptides. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background). The average and standard deviation (error bars) of five mice (n=5) are shown for each group. Splenocytes were isolated 14 days after the vaccination.

[0020] **FIG. 9** shows a schematic diagram of an exemplary pan-influenza vaccine. In brief, a fusion protein comprising a nucleoprotein from an Influenza Type A virus (FluA) and a nucleoprotein from an Influenza Type B virus (FluB) is expressed from mRNA, self-replicating RNA, or temperature-sensitive, self-replicating RNA (srRNAts) delivered to mammalian host cells. In exemplary embodiments, the coding region of the fusion protein is the gene of interest (GOI) inserted within the srRNAts. Specifically, G5010 is a fusion protein comprising the signal peptide sequence from the human CD5 antigen (CD5-SP) set forth as SEQ ID NO:8, the FluA nucleoprotein (Influenza Type A, H5N8 subtype [A/breeder duck/Korea/Gochang1/2014], GenBank No. KJ413835.1, ProteinID No. AHL21420.1), and the FluB nucleoprotein (Influenza Type B [B/Florida/4/2006], GenBank No. CY033879.1, ProteinID No. ACF54251.1). In G5010, the CD5-SP replaces the start methionine of the FluA nucleoprotein, and the FluA nucleoprotein is fused to the methionine of the start codon of the FluB nucleoprotein.

[0021] **FIG. 10** shows an alignment of the nucleoprotein of Influenza A (H5N8 strain; ProteinID AHL21420.1) used as a vaccine antigen in **G5010** (SEQ ID NO:13) and the nucleoprotein of Influenza A (NP/AnnArbor H2N2; ProteinID P21433) used as a source (SEQ ID NO:17) of a peptide pool for ELISpot assay.

[0022] **FIG. 11A** and **FIG. 11B** show the frequency of cytokine-secreting cells in samples of splenocytes obtained from BALB/c mice that had been immunized by a single intradermal injection of 100 µL solution containing either 5 µg or 25 µg of a temperature-controllable self-replicating RNA (srRNA1ts2 [PCT/US20/67506]) encoding the **G5010** antigen or a placebo (PBO: buffer only). **FIG 11A** shows the frequency of interferon-gamma (INF-γ) spot-forming cells (SFC) and **FIG. 11B** shows the frequency of interleukin-4 (IL-4) SFC in 1x10^6 splenocytes after restimulation by culturing the splenocytes in the presence or absence of a pool of Influenza A (H2N2) nucleoprotein peptides. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of

peptides (background). The average and standard deviation (error bars) of five mice (n=5) are shown for each group. Splenocytes were isolated 14 days after the vaccination.

[0023] **FIG. 12** shows a schematic diagram of an exemplary pan-ebolavirus vaccine. In brief, a fusion protein comprising nucleoproteins of four different ebolavirus strains is expressed from mRNA, self-replicating RNA, or temperature-sensitive, self-replicating RNA (srRNAts) delivered to mammalian host cells. In exemplary embodiments, the coding region of the fusion protein is the gene of interest (GOI) inserted within the srRNAts. Specifically, the exemplary PanEbola antigen is a fusion protein comprising the signal peptide sequence from the human CD5 antigen (CD5-SP) set forth as SEQ ID NO:8, a part of a nucleoprotein of Zaire ebolavirus (residues 2-739; total 738 aa; GenBank ID: AF272001) set forth as SEQ ID NO:18, a part of a nucleoprotein of Sudan ebolavirus (residues 403-738; total 336 aa; GenBank ID: AF173836) set forth as SEQ ID NO:19, a part of a nucleoprotein of Bundibugyo ebolavirus (residues 403-739; total 337 aa; GenBank ID: FJ217161) set forth as SEQ ID NO:20, and a part of a nucleoprotein of Taï Forest ebolavirus (residues 483-651; total 169 aa; GenBank ID: FJ217162) set forth as SEQ ID NO:21. The amino acid sequence of the PanEbola antigen is set forth as SEQ ID NO:22, while the nucleic acid sequence encoding the PanEbola antigen is set forth as SEQ ID NO:23.

[0024] **FIG. 13** shows amino acid sequence similarities among four species of Ebolavirus as percent identities. Amino acid sequence of Zaire ebolavirus NP (GenBank ID: AF272001), Sudan ebolavirus NP (GenBank ID: AF173836), Bundibugyo ebolavirus NP (GenBank ID: FJ217161), Taï Forest ebolavirus NP (GenBank ID: FJ217162) were compared to each other by using NCBI BlastP algorithm. Based on the sequence alignment, proteins were divided into well-conserved regions (A) and less well-conserved regions (B). The amino acid sequence identity between Zaire ebolavirus NP and Sudan ebolavirus NP was 88% for Region A, whereas it was 42% for Region B. The amino acid sequence identity between Zaire ebolavirus NP and Bundibugyo ebolavirus NP was 92% for Region A, whereas it was 53% for Region B. The amino acid sequence identity between Zaire ebolavirus NP and Taï Forest ebolavirus NP was 92% for Region A, whereas it was 54% for Region B. For Region B, Bundibugyo (B) and Taï Forest (B) sequences shared a relatively high level of sequence similarity. Based on the sequence alignment of Region B, proteins were divided into well-conserved regions (80% and 86%

similarity; no label) and a less well-conserved region (40% identity; referred to herein as Region C).

[0025] **FIG. 14A** and **FIG. 14B** show the frequency of cytokine-secreting cells in samples of splenocytes obtained from BALB/c mice that had been immunized by a single intradermal injection of 100 µL solution containing either 25 µg of a temperature-controllable self-replicating RNA (srRNA1ts2 as described in WO 2021/138447 A1, also called c-srRNA) encoding the PanEbola antigen (srRNA1ts2-PanEbola, also called G5011) or a placebo (PBO: buffer only). **FIG 14A** shows the frequency of interferon-gamma (INF-γ) spot-forming cells (SFC) and **FIG. 14B** shows the frequency of interleukin-4 (IL-4) SFC in 1x10^6 splenocytes after restimulation by culturing the splenocytes in the presence or absence of a pool of 182 peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein (Swiss-Prot ID: B8XCN6) of Taï Forest Ebolavirus [JPT peptide; PepMix Tai Forest Ebolavirus (NP); JPT Product Code: PM-TEBOV-NP]. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background). The average and standard deviation (error bars) of five mice (n=5) are shown for each group. Splenocytes were isolated 14 days after the vaccination.

[0026] **FIG. 15** depicts a schematic diagram showing exemplary srRNA1ts2 constructs encoding the receptor binding domain (RBD) of the spike protein of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). G5003 is the same antigen as "srRNA1ts2-2019-nCoV-RBD1" presented in FIG. 21 of WO 2021/138447 A1; and G5003 encodes a fusion protein including the signal peptide of CD5 (residues 1-24) and the RBD of the spike protein of SARS-CoV-2 (an original Wuhan strain). **G5003o** encodes a fusion protein (SEQ ID NO:25) including the signal peptide of CD5 (residues 1-24) and the RBD of the spike protein of SARS-CoV-2 (an omicron strain B.1.1.529: Science Brief: Omicron (B.1.1.529) Variant | CDC).The nucleotide sequence of the G5003o open reading frame is set forth as SEQ ID NO:24.

[0027] **FIG. 16A** and **FIG. 16B** show the frequency of cytokine-secreting cells in samples of splenocytes obtained from C57BL/6 mice that had been immunized by a single intradermal injection of 100 µL solution containing either placebo (PBO: buffer only) or 25 µg of a temperature-controllable self-replicating RNA (srRNA1ts2 as described in WO 2021/138447 A1) encoding the G5003o antigen. **FIG 16A** shows the frequency of

interferon-gamma (INF-γ) spot-forming cells (SFC) and **FIG. 16B** shows the frequency of interleukin-4 (IL-4) SFC in 1x10^6 splenocytes from immunized mice restimulated by culturing in the presence or absence of a pool of 53 peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through RBD of SARS-CoV-2 omicron variant (B.1.1.529) [JPT peptide Product Code: PM-SARS2-RBDMUT08-1]. The assays were performed by the ELISpot assay. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background). The average and standard deviation (error bars) of five mice (n=5) are shown for each group. Splenocytes were isolated 14 days after the vaccination.

[0028] **FIG. 17A-17C** show the induction of both cellular immunity and humoral immunity in mice as a consequence of administering a composition comprising a c-srRNA encoding an antigen, followed by administering a composition comprising a protein antigen. **FIG. 17A** depicts a schematic diagram of experimental procedures. On day -40, blood was withdrawn from female BALB/c mice for the plaque reduction neutralization test (PRNT). On day -36, these mice were treated with c-srRNA encoding G5003 antigen. The c-srRNA was injected intradermally into mouse skin as a naked RNA, without any nanoparticle nor transfection reagent. On day -22 (14 days after c-srRNA-G5003 vaccination), a half of mice were sacrificed to obtain splenocytes for ELISpot assays. On day 0, the remaining mice were intradermally injected with a Spike protein of SARS-CoV-2 Delta variant (B.1.617.2) mixed with adjuvant (AddaVax™ adjuvant marketed by Invivogen). On day 7 (7 days after the Spike protein injection), blood was withdrawn for the PRNT assays. **FIG. 17B** shows the induction of cellular immunity against the RBD protein by a single intradermal vaccination with the c-srRNA-G5003 vaccine. The figure shows the frequency of interferon-gamma (INF-γ) spot-forming cells (SFC) in 1x10^6 splenocytes from immunized mice restimulated by culturing in the presence or absence of a pool of 53 peptides (15mers with 11 amino acid overlaps) that covers SARS-CoV-2 RBD (an original Wuhan strain). The assays were performed by the ELISpot assay. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background). The average and standard deviation (error bars) of five mice (n=5) are shown for each group. Splenocytes were isolated on day -22 (14 days after the vaccination). **FIG. 17C** shows the titer of serum antibodies that can neutralize (50%) the SARS-CoV-2 virus (Delta variant B.1.617.2), measured by a

plaque reduction neutralization assay (PRNT). Exposure to a spike protein of SARS-CoV-2 virus (Delta variant B.1.617.2) induced neutralization antibodies specifically against the Delta variant of SARS-CoV-2 virus only in mice vaccinated with a vaccine c-srRNA-G5003, encoding the RBD of the SARS-CoV-2 (an original Wuhan strain).

[0029] **FIG. 18A--18C** show the induction of cellular immunity in mice as a consequence of administering a composition comprising a protein antigen, followed by administering a composition comprising c-srRNA encoding an antigen. **FIG. 18A** depicts a schematic diagram of experimental procedures. On day 0 (1st treatment), female C57BL/6 mice were treated with intradermal injection with 10 μg RBD protein (Sino Biological SARS-CoV-2 [2019-nCoV]) + Adjuvant (AddaVax™ adjuvant marketed by Invivogen). On day 14 (2nd treatment), the mice were treated with intradermal injection of a placebo (PBO: buffer only), 25 μg c-srRNA encoding G5003 antigen, 25 μg c-srRNA encoding G5003o antigen, or 10 μg RBD protein (Sino Biological SARS-CoV-2 [2019-nCoV]) + Adjuvant (AddaVax™ adjuvant). On day 28, mice were sacrificed, and splenocytes and serum were collected. **FIG. 18B** shows the frequency of interferon-gamma (INF-γ) and **FIG. 18C** shows the frequency of interleukin 4 (IL-4) spot-forming cells (SFC) in 1x10^6 splenocytes restimulated by culturing in the presence or absence of a pool of 53 peptides (15mers with 11 amino acid overlaps) that covers SARS-CoV-2 RBD (an original Wuhan strain). The assays were performed by the ELISpot assay. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background).

[0030] **FIG. 19** shows the level of serum antibodies against the RBD of the SARS-CoV-2 virus (an original Wuhan strain) as determined by an ELISA assay (represented by the OD450 measurement). The average and standard deviation (error bars) of five mice (n=5) are shown for each group. The data of Day -1 (before the 1st treatment) and the data of Day 28 (after the 2nd treatment) are shown for each group.

[0031] **FIGS. 20A-D** show the frequency of interferon-gamma (INF-γ)- or interleukin 4 (IL-4)-secreting cells in samples of splenocytes obtained from BALB/c mice that had been immunized by a single intradermal injection of 100 μL solution containing either 5 μg (n=1) or 25 μg (n=4) of a temperature-controllable self-replicating RNA (srRNA1ts2 as described in WO 2021/138447 A1) encoding the G5006 antigen (FIG. 2) or a placebo (PBO: buffer only:

n=5). The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background). The average and standard deviation (error bars) of one mouse (n=1) or four mice (n=4) are shown for each group. Splenocytes were isolated 14 days after the vaccination. **FIG. 20A** and **FIG. 20B** show the results after restimulation by culturing the splenocytes in the presence or absence of a pool of SARS-CoV-2 nucleoprotein peptides. **FIG. 20C** and **FIG. 20D** show the results after restimulation by culturing the splenocytes in the presence or absence of a pool of MERS-CoV-2 nucleoprotein peptides.

[0032] **FIG. 21** shows survival (%) of the female BALB/c mice vaccinated with c-srRNA-G5006, followed by the injection of tumor cells expressing G5006 antigens.

[0033] **FIG. 22** depicts a schematic diagram showing exemplary srRNA1ts2 constructs encoding a fusion protein of the signal peptide of CD5 (residues 1-24), the receptor binding domain (RBD) of the spike protein of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), the nucleoprotein of SARS-CoV-2, the nucleoprotein of MERS-CoV, and the RBD of MERS-CoV (named here G5006d). The amino acid sequence of the pan-coronavirus antigen (G5006d) is set forth as SEQ ID NO:27, and the nucleotide sequence of its open reading frame is set forth as SEQ ID NO:26.

[0034] **FIGS. 23A-B** show the frequency of cytokine-secreting cells in samples of splenocytes obtained from female C57BL/6 mice that had been immunized by a single intradermal injection of 100 μL solution containing either placebo (PBO: buffer only), 25 μg of a temperature-controllable self-replicating RNA (srRNA1ts2 as described in WO 2021/138447 A1) encoding the G5006 antigen, or 25 μg of a temperature-controllable self-replicating RNA (srRNA1ts2 as described in WO 2021/138447 A1) encoding the G5006d antigen. **FIG. 23A** shows the frequency of interferon-gamma (INF-γ) spot-forming cells (SFC) and **FIG. 23B** shows the frequency of interleukin-4 (IL-4) SFC in 1x10^6 splenocytes from immunized mice restimulated by culturing in the presence or absence of pools of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through (A) RBD of Spike protein of SARS-CoV-2 [JPT Peptide Product Code: PM-WCPV-S-RBD-2]; (B) Nucleoprotein of SARS-CoV-2 [JPT peptide Product Code: PM-WCPV-NCAP]; (C) Nucleoprotein of MERS-CoV [JPT peptide, custom made]; and (D) Spike protein of MERS-CoV [JPT peptide Product Code: PM-MERS-

CoV-S-1). The assays were performed by the ELISpot assay. The frequency obtained in the presence of peptides is plotted in the graph after subtracting the frequency obtained in the absence of peptides (background). The average and standard deviation (error bars) of five mice (n=5) for PBO, four mice (n=4) for G5006, and five mice (n=5) for G5006d, are shown for each group. Splenocytes were isolated 14 days after the vaccination.

[0035] **FIG. 24** depicts a schematic diagram showing exemplary srRNA1ts2 constructs encoding a fusion protein (G5012) of the signal peptide of CD5 (residues 1-24), a part of the hemagglutinin (HA) of the Influenza A (A/New Caledonia/20/1999(H1N1)) (residues 25-165), nucleoprotein of Influenza A (A/breeder duck/Korea/Gochang1/2014(H5N8)) (residues 166-662), nucleoprotein of Influenza B (B/Florida/4/2006) (residues 663-1222), and a part of the hemagglutinin (HA) of the Influenza B (B/Florida/4/2006) (residues 1223-1365). The amino acid sequence of the pan-influenza virus antigen (G5012) is set forth as SEQ ID NO:29, and the nucleotide sequence of its open reading frame is set forth as SEQ ID NO:28.

[0036] **FIG. 25** shows the effects of Chitosan Oligomers on gene (luciferase) expression from srRNA1ts2 (exemplary c-srRNA) in mice. c-srRNA encoding luciferase was intradermally injected into mice under the following conditions: 1, a control – c-srRNA only; 2, c-srRNA mixed with chitosan oligosaccharide (0.001 μg/mL); 3, c-srRNA mixed with chitosan oligosaccharide (0.01 μg/mL); 4, c-srRNA mixed with chitosan oligosaccharide (0.5 μg/mL); and 5, c-srRNA mixed with chitosan oligosaccharide lactate (0.1 μg/mL).

DETAILED DESCRIPTION

[0037] Broader, longer-lasting protection against SARS-CoV-1, SARS-CoV-2, MERS-CoV, and their variants, is best achieved through vaccines that induce cellular immunity (i.e., T-cell-inducing vaccines involving CD8+ killer T cells and CD4+ helper T cells). This is a departure from the current, neutralizing antibody-focused COVID-19 vaccine paradigm, as discussed in the Background section. The critical importance of cellular immunity in fighting against coronaviruses has been demonstrated experimentally and extensively discussed [Sette and Crotty 2021]. Cellular immunity alone can provide protection via CD8+ killer T cells [Matchett et al., 2021]. Also, cellular immunity depends on linear T cell epitopes, whereas humoral immunity depends on conformational (as well as linear) B cell epitopes. Therefore, cellular immunity is much more robust against variants than humoral immunity. Furthermore,

memory T cells last longer than memory B cells, and thus, potentially provide lifelong immunity. This requires both suitable antigens and a cellular immunity-based vaccine platform.

Cellular immunity-based mRNA vaccine platform

[0038] The vaccine platform is described in Elixirgen's earlier patent application [PCT/US20/67506, now published as WO 2021/138447 A1]. This vaccine platform is optimized to induce cellular immunity, which becomes possible by combining existing knowledge of vaccine biology with temperature-controllable self-replicating mRNA (srRNAts) based on an Alphavirus, such as the Venezuelan equine encephalitis virus (VEEV). The terms c-srRNA and srRNAts are used interchangeably throughout the present disclosure, with srRNA1ts2 (described in WO 2021/138447 A1) being an exemplary embodiment. srRNAts is based on srRNA, also known as self-amplifying mRNA (saRNA or SAM), by incorporation of small amino acid changes in the Alphavirus replicase that provide temperature-sensitivity. Elixirgen Therapeutic Inc.'s srRNAts is functional at 30-35°C, but not functional at or above 37°C ± 0.5°C. It carries all the benefits of mRNA platforms: no genome integration, rapid development and deployment, and a simple good manufacturing process (GMP), as well as additional advantages of srRNA platforms compared to mRNA platforms, particularly longer expression [Johanning et al., 1995] and higher immunogenicity at a lower dosage [Brito et al., 2014]. However, this simple temperature-controllable feature makes it possible to pull together many desirable features of T-cell inducing vaccine as described herein.

[0039] In brief, srRNA1ts2 is a temperature-sensitive, self-replicating VEEV-based RNA replicon developed for transient expression of a heterologous protein. Temperature-sensitivity is conferred by an insertion of five amino acids residues within the non-structural Protein 2 (nsP2) of VEEV. The nsP2 protein is a helicase/proteinase, which along with nsP1, nsP3 and nsP4 constitutes a VEEV replicase. srRNA1ts2 does not contain VEEV structural proteins (capsid, E1, E2 and E3). The disclosure of WO 2021/138447 A1 of Elixirgen Therapeutics, Inc. is hereby incorporated by reference. In particular, Example 3, Figure 12, and SEQ ID NOs. 29-49 of WO 2021/138447 A1 are hereby incorporated by reference.

[0040] Overall, the srRNAts platform's compelling potential for immunogenicity (dose-sparing) and safety benefits (temperature-control and naked delivery), provisioning of long-lasting baseline cellular immunity, and ability to provide rapid humoral responses across variants

makes it a strong candidate for large-scale deployment to meet the global need for an inexpensive, safe, variant-addressing vaccine that provides long-term immunity.

General Techniques and Definitions

[0041] The practice of the present disclosure will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry and immunology, which are within the skill of the art.

[0042] As used herein and in the appended claims, the singular forms "a", "an", and "the" include plural references unless indicated otherwise. For example, "an" excipient includes one or more excipients.

[0043] The phrase "comprising" as used herein is open-ended, indicating that such embodiments may include additional elements. In contrast, the phrase "consisting of" is closed, indicating that such embodiments do not include additional elements (except for trace impurities). The phrase "consisting essentially of" is partially closed, indicating that such embodiments may further comprise elements that do not materially change the basic characteristics of such embodiments.

[0044] The term "about" as used herein in reference to a value, encompasses from 90% to 110% of that value (e.g., molecular weight of about 5,000 daltons when used in reference to a chitosan oligosaccharide refers to 4,500 daltons to 5,500 daltons).

[0045] The term "antigen" refers to a substance that is recognized and bound specifically by an antibody or by a T cell antigen receptor. Antigens can include peptides, polypeptides, proteins, glycoproteins, polysaccharides, complex carbohydrates, sugars, gangliosides, lipids and phospholipids; portions thereof and combinations thereof. In the context of the present disclosure, the term "antigen" typically refers to a polypeptide or protein antigen at least eight amino acid residues in length, which may comprise one or more post-translational modifications.

[0046] The terms "polypeptide" and "protein" are used interchangeably to refer to a polymer of amino acid residues, and are not limited to a certain length unless otherwise specified. Polypeptides may include natural amino acid residues or a combination of natural and non-natural amino acid residues. The terms also include post-expression modifications of the polypeptide, for example, glycosylation, sialylation, acetylation, phosphorylation, and the like. In

some aspects, the polypeptides may contain modifications with respect to a native or natural sequence, as long as the protein maintains the desired activity (e.g., antigenicity).

[0047] The terms "isolated" and "purified" as used herein refers to a material that is removed from at least one component with which it is naturally associated (e.g., removed from its original environment). The term "isolated," when used in reference to a recombinant protein, refers to a protein that has been removed from the culture medium of the host cell that produced the protein. In some embodiments, an isolated protein (e.g., SARS-CoV-2 Spike protein) is at least 75%, 90%, 95%, 96%, 97%, 98% or 99% pure as determined by HPLC.

[0048] An "effective amount" or a "sufficient amount" of a substance is that amount sufficient to effect beneficial or desired results, including clinical results, and, as such, an "effective amount" depends upon the context in which it is being applied. In the context of administering a composition of the present disclosure comprising an mRNA encoding an antigen, an effective amount contains sufficient mRNA to stimulate an immune response (preferably a cellular immune response against the antigen).

[0049] In the present disclosure, the terms "individual" and "subject" refer to a mammals. "Mammals" include, but are not limited to, humans, non-human primates (e.g., monkeys), farm animals, sport animals, rodents (e.g., mice and rats) and pets (e.g., dogs and cats). In some preferred embodiments, the subject is a human subject.

[0050] The term "dose" as used herein in reference to a composition comprising a mRNA encoding an antigen refers to a measured portion of the taken by (administered to or received by) a subject at any one time. Administering a composition of the present disclosure to a subject in need thereof, comprises administering an effective amount of a composition comprising a mRNA encoding an antigen to stimulate an immune response to the antigen in the subject.

[0051] "Stimulation" of a response or parameter includes eliciting and/or enhancing that response or parameter when compared to otherwise same conditions except for a parameter of interest, or alternatively, as compared to another condition (e.g., increase in antigen-specific cytokine secretion after administration of a composition comprising or encoding the antigen as compared to administration of a control composition not comprising or encoding the antigen).

For example, "stimulation" of an immune response (e.g., Th1 response) means an increase in the response. Depending upon the parameter measured, the increase may be from 2-fold to 200-fold or over, from 5-fold to 500-fold or over, from 10-fold to 1000-fold or over, or from 2, 5, 10, 50, or 100-fold to 200, 500, 1,000, 5,000, or 10,000-fold.

[0052] Conversely, "inhibition" of a response or parameter includes reducing and/or repressing that response or parameter when compared to otherwise same conditions except for a parameter of interest, or alternatively, as compared to another condition. For example, "inhibition" of an immune response (e.g., Th2 response) means a decrease in the response. Depending upon the parameter measured, the decrease may be from 2-fold to 200-fold, from 5-fold to 500-fold or over, from 10-fold to 1000-fold or over, or from 2, 5, 10, 50, or 100-fold to 200, 500, 1,000, 2,000, 5,000, or 10,000-fold.

[0053] The relative terms "higher" and "lower" refer to a measurable increase or decrease, respectively, in a response or parameter when compared to otherwise same conditions except for a parameter of interest, or alternatively, as compared to another condition. For instance, a "higher antibody titer" refers to an antigen-reactive antibody titer as a consequence of administration of a composition of the present disclosure comprising an mRNA encoding an antigen that is at least 2, 3, 4, 5, 6, 7, 8, 9, or 10-fold above an antigen-reactive antibody titer as a consequence of a control condition (e.g., administration of a comparator composition that does not comprise the mRNA or comprises a control mRNA that does not encode the antigen). Likewise, a "lower antibody titer" refers to an antigen-reactive antibody titer as a consequence of a control condition (e.g., administration of a comparator composition that does not comprise the mRNA or comprises a control mRNA that does not encode the antigen) that is at least 2, 3, 4, 5, 6, 7, 8, 9, or 10-fold below an antigen-reactive antibody titer as a consequence of administration of a composition of the present disclosure comprising an mRNA encoding an antigen.

[0054] As used herein the term "immunization" refers to a process that increases a mammalian subject's reaction to antigen and therefore improves its ability to resist or overcome infection and/or resist disease.

[0055] The term "vaccination" as used herein refers to the introduction of a vaccine into a body of a mammalian subject.

[0056] As used herein, "percent (%) amino acid sequence identity" and "percent identity" and "sequence identity" when used with respect to an amino acid sequence (reference polypeptide sequence) is defined as the percentage of amino acid residues in a candidate sequence (e.g., the subject antigen) that are identical with the amino acid residues in the reference polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for aligning sequences, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

[0057] An amino acid substitution may include replacement of one amino acid in a polypeptide with another amino acid. Amino acid substitutions may be introduced into an antigen of interest and the products screened for a desired activity, e.g., increased stability and/or immunogenicity.

[0058] Amino acids generally can be grouped according to the following common side-chain properties:

(1) hydrophobic: Norleucine, Met, Ala, Val, Leu, Ile;

(2) neutral hydrophilic: Cys, Ser, Thr, Asn, Gln;

(3) acidic: Asp, Glu;

(4) basic: His, Lys, Arg;

(5) residues that influence chain orientation: Gly, Pro; and

(6) aromatic: Trp, Tyr, Phe.

[0059] Conservative amino acid substitutions will involve exchanging a member of one of these classes with another member of the same class. Non-conservative amino acid substitutions will involve exchanging a member of one of these classes with a member of another class.

[0060] As used herein, the term "excipient" refers to a compound present in a composition comprising an active ingredient (e.g., mRNA encoding an antigen). Pharmaceutically acceptable excipients are inert pharmaceutical compounds, and may include for instance, solvents, bulking agents, buffering agents, tonicity adjusting agents, and preservatives (Pramanick et al., Pharma Times, 45:65-77, 2013). In some embodiments the compositions of the present disclosure comprise an excipient that functions as one or more of a solvent, a bulking agent, a buffering agent, and a tonicity adjusting agent (e.g., sodium chloride in saline may serve as both an aqueous vehicle and a tonicity adjusting agent).

Optimized for intradermal delivery

[0061] Intradermal vaccination results in long-lasting cellular immunity and increased immunogenicity [Hickling and Jones, 2009]. Human skin (epidermis and dermis) is rich in antigen-presenting cells (APCs), including Langerhans cells and dermal dendritic cells (DCs). Intradermal vaccination is known to be 5- to 10-times more effective than subcutaneous or intramuscular vaccination because it targets the APCs present in skin [Hickling and Jones, 2009], thereby activating the T cell immunity pathway for long-lasting immunity. By intradermal injection, srRNAts is predominantly taken up by skin APCs, wherein it replicates, produces antigen, digests the antigen into peptides, and presents the peptides to T cells (**FIG. 1**). The peptides presented through this pathway stimulates MHC-I-restricted CD8+ killer T cells. In an alternative pathway, APCs also take antigens produced by nearby skin cells. The peptides presented through this pathway stimulate MHC-II-restricted CD4+ Helper T cells, which helps B cells to produce neutralizing antibodies (nAb) to fight virus infection.

Issues and solutions for intradermal injection

[0062] Here are potential issues that we have identified and the solutions that the srRNAts platform offers.

[0063] (1) A key unrecognized hurdle for the application of srRNA as an intradermal vaccine platform is that both mRNA and srRNA do not express antigen well at skin temperature [PCT/US20/67506]. Unintuitively, the temperature of the human skin is lower (about 30-35°C) than human core body temperature (about 37°C); this means that vectors and platforms developed at 37°C are not optimal for intradermal injection. One innovation of the srRNAts

platform is that it expresses antigen strongly at skin temperature [PCT/US20/67506]. Furthermore, this temperature-control also minimizes the safety risk caused by unintended systemic distribution of srRNAts because srRNAts becomes inactivated once its temperature increases above its permissive threshold (when it moves closer to the core of the body). In other words, the srRNAts platform expresses antigen the best for intradermal injection compared to mRNA and srRNA, and it additionally has safety features: the vector's ability to spread and become produced in other areas of a subject's body is limited or inactivated.

[0064] (2) Another challenge for intradermal vaccination is the lack of suitable additives. Because adjuvants such as aluminum-salt and oil-in-water are too reactogenic locally when delivered by the intradermal route, no adjuvant has been incorporated into clinically approved intradermal vaccines, resulting in lower immunogenicity [Hickling and Jones, 2009]. Lipid Nanoparticles (LNPs) used for mRNA and srRNA vaccines, which are administered intramuscularly, are also oil-in-water, which may cause skin reactogenicity and increase risk of allergic reactions to LNP components such as PEG. The c-srRNA platform is a solution to this problem since it is injected as naked c-srRNA (no LNPs, no adjuvants). First, self-replication of RNAs inside cells, especially APCs, induces the strong innate immunity, which substitutes the major functions of adjuvants. Second, data in the literature and obtained during development of the present disclosure demonstrates that, specifically for intradermal injection, naked mRNA/srRNA is equally efficient to produce an antigen compared to electroporation of mRNA/srRNA [Johansson et al, 2012] and mRNA/srRNAs combined with LNPs [Golombek et al., 2018].

[0065] (3) A third challenge is the limited number of precedents for intradermal vaccines. Only the BCG vaccine has been administered intradermally on a routine basis, and currently available COVID-19 vaccines are all administered intramuscularly. One way we lower the hurdle for adopting intradermal injection is by using specialized devices such as the MicronJet600 (NanoPass) and Immucise (Terumo), which are now available to enable easy, consistent intradermal injection. These devices are also good candidates for large-scale production and deployment. However, due to a relatively high cost of these special devices, an intradermal injection by the Mantoux technique using a standard needle and syringe is also an option.

Design of suitable antigens

[0066] The cellular immunity-focused approach allowed for the reconsideration of all the proteins encoded on viral genomes as antigen candidates, as humoral immunity, i.e., the induction of neutralizing antibodies, is not the primary consideration.

[0067] When selecting an antigen that would provide broader protection against SARS-CoV-1, SARS-CoV-2, MERS-CoV, and their variants, the Nucleoprotein (N) was determined to be the most suitable, because (1) N is the most abundant protein, followed by Membrane (M) and Spike (S) in viral particles [Finkel et al., 2021], (2) N is overall the most conserved protein among the above indicated Betacoronaviruses [Grifoni et al., 2020], and (3) epitopes for B and T cells are the most abundant in S and N [Grifoni et al., 2020]. This is consistent with the earlier proposal that N is the best antigen for the vaccine [Dutta et al., 2020]. Notably, a recent report clearly demonstrated that a vaccine using N alone as an antigen can provide an S-independent protective immunity in both hamster and mouse [Matchett et al., 2021]. Although disease enhancement was observed for N vaccines, as well as S vaccines previously [Lambert et al., 2020], these data were obtained by using different vectors with unfavorable Th2>Th1 profiles.

[0068] An exemplary vaccine candidate, srRNA1ts2-G5005, was designed to express the N protein of SARS-CoV-2 (SARS2-N). However, MERS-N forms a distinct group and shows only 48% identity [Tilocca et al., 2020]. With this in mind, a further exemplary vaccine candidate, srRNA1ts2-G5006, was designed to express a fusion protein of SARS2-N and MERS-N. The G5005 and G5006 antigens are shown schematically in **FIG. 2**. srRNA1ts2-G5005 is suitable for induction of immune responses against SARS-CoV-1, SARS-CoV-2, and their variants. In contrast, srRNA1ts2-G5006 is suitable for induction of a pan-coronavirus immune response (e.g., against SARS-CoV-1, SARS-CoV-2, MERS-CoV, and their variants).

[0069] To address the emergence of a variant (mutated) form of SARS-CoV-2 virus, c-srRNA encoding the RBD of SARS-CoV-2 omicron variant (G5003o) was generated and intradermally administered to C57BL/6 mice (Example 8 and **FIG. 15**). Cellular immunity was assessed 14 days after the vaccination. The results clearly demonstrate that c-srRNA can induce omicron variant-specific cellular immunity, when the open reading frame of the receptor binding domain (RBD) of the omicron variant is included in the c-srRNA. Importantly, c-srRNA

encoding the G5003o antigen was found to induce a Th1-biased response as shown in **FIG. 16A-16B** [Th1 (INF-γ) > Th2 (IL-4], which is favored for vaccines.

Inclusion in prime-boost immunization regimens

[0070] One of the unique features of intradermally administered c-srRNA vaccine is its ability to induce cellular immunity without apparent induction of humoral immunity (i.e., antibodies). As determined during development of the present disclosure, c-srRNA vaccines are able to prime a humoral immune response to a subsequently encountered protein antigen. In brief, mice were first treated with c-srRNA encoding an antigen (i.e., RBD of SARS-CoV-2 Wuhan strain) and were subsequently treated with an adjuvanted variant RBD protein (i.e., RBD of SARS-CoV-2 Delta variant) as described in Example 9 and shown in **FIG. 17A**.

[0071] Cellular immunity, assessed by measuring the presence of antigen-specific IFN-γ-secreting T cells, was already induced by day 14 post-primary vaccination (prime) as shown in **FIG. 17B**. Antigen-specific antibody was not detected at this time. After treatment with the adjuvanted protein antigen, antibodies were induced as early as day 7 post-secondary vaccination (boost) as shown in **FIG. 17C**. This early induction of antibodies is consistent with a secondary immune response, indicating that c-srRNA already primed humoral immunity. Importantly, the antibodies induced by the protein antigen boost were able to neutralize the viral variant, which has a distinct RBD sequence from the RBD antigen encoded by the c-srRNA vaccine. This surprising finding indicates that the c-srRNA vaccine can induce a protective immune response against a pathogen with an antigen sequence that differs from the antigen sequence encoded by the c-srRNA vaccine. Thus, the c-srRNA vaccines are expected to induce broadly reactive immune responses, which are critical for providing protection against variant pathogens.

[0072] Subunit vaccines against pathogens generally do not provide the long-lasting humoral immunity (i.e., pathogen-specific antibodies), and therefore one or more booster vaccines are required. As determined during development of the present disclosure, c-srRNA vaccines are suitable for use as a booster vaccine, when an adjuvanted protein is administered as a prime vaccine. In brief, mice were first treated with adjuvanted protein (i.e., RBD of SARS-CoV-2 Wuhan strain) and were subsequently treated with a placebo (PBO: buffer only), c-srRNA encoding G5003 antigen (Wuhan RBD), c-srRNA encoding G5003o antigen (Omicron

RBD), or the adjuvanted protein antigen (Wuhan RBD) as described in Example 10 and shown in **FIG. 18A**.

[0073] As shown in **FIG. 17C**, c-srRNA vaccine alone does not induce humoral immunity in the form of a neutralizing antibody response (see, PBO day 7). However, when humoral immunity is primed by the adjuvanted protein (as a model for primary vaccination), the c-srRNA booster vaccine is able to induce both antigen-specific cytokine responses (**FIG. 18B-18C**) and antigen-specific antibody responses (**FIG. 19**). It is worth noting that in the current experimental condition, a single dose of adjuvanted protein did not induce RBD-specific antibodies. Apparently, cellular immunity induced by c-srRNA is capable of stimulating antibody production to an earlier encountered protein antigen. This observation is indicative of important interactions occurring between cellular and humoral immune responses.

Elimination of antigen-expressing cells in vivo

[0074] c-srRNA vaccines are able to induce strong cellular immune responses (i.e., antigen-specific CD8+ cytotoxic T lymphocytes and CD4+ helper T lymphocytes). Antigen-specific CD8+ CTL lyse cells in which the antigen is expressed. Antigen recognition by CD8+ CTL is based on presentation of short peptide fragments (T cell epitopes) by MHC class I molecules, and thus, the antigen does not have to be expressed on the surface of target cells. For a vaccine directed against a pathogen, the vaccine is expected to lyse cells infected with the pathogen. For a vaccine directed against a cancer, the vaccine is expected to lyse cancer cells.

[0075] A c-srRNA vaccine encoding a fusion protein of SARS-CoV-2 nucleoprotein and MERS-CoV nucleoprotein (called SMN protein or G5006) as an antigen was produced. In order to model cells infected with a virus, a 4T1 breast cancer cell line derived from BALB/c mouse and known as a model for a triple-negative stage IV human breast cancer was selected. When injected into BALB/c mice, the 4T1 cells grow rapidly and form tumors. This syngeneic mouse model was used to mimic the rapid increase of infected cells. The 4T1 cells expressing the SMN protein (named 4T1-SMN) was established by transfecting a plasmid vector encoding an SMN protein under the CMV promoter, so that the protein is constitutively expressed in 4T1 cells. The fusion protein is the same as G5006 except that the CD5 signal peptide was removed from the N-terminus of the SMN protein expressed in 4T1 cells.

[0076] BALB/c mice were vaccinated with c-srRNA-G5006, and the induction of cellular immunity was demonstrated by the presence of T-cells that responded to both SARS-CoV-2 nucleoprotein (**FIG. 20A-20B**) and MERS-CoV nucleoprotein (**FIG. 20C-20D**). Subsequently, 4T1-SMN cells were injected into the BALB/c mice vaccinated with c-srRNA-G5006 on day 24 (24 days post-vaccination). As expected, 4T1-SMN cells grew rapidly in mice that received a placebo (no vaccine group). In contrast, the growth of 4T1-SMN tumors was suppressed in the c-srRNA-G5006 vaccinated mice. In two mice that received 25 μg of the c-srRNA-G5006 vaccine, while the tumors initially grew, the mice eventually became tumor-free and survived long after the recipients of the placebo had died. Furthermore, even after the second round of injection of 4T1-SMN tumors on day 143 after vaccination, no tumors grew, and the mice remained alive and tumor-free for the duration of the study (**FIG. 21**). This result suggests that a c-srRNA vaccine encoding the G5006 antigen (i.e., SMN protein) can induce a protective immune response by elimination of cells infected with SARS-CoV-2 or MERS-CoV.

Pancoronavirus booster vaccine

[0077] For infectious diseases, such as COVID-19, World Health Organization guidelines require a licensed vaccine to be capable of inducing neutralizing antibodies (nAb). This requirement makes sense since nAb can prevent cells from becoming infected, and thus nAb can efficiently control the spread of infection. However, nAb levels generally decline rapidly, and therefore booster vaccines are needed periodically (e.g., once or twice a year) after completion of a primary vaccination series (1st and 2nd vaccinations) to maintain adequate nAb levels. The high mutation rate of SARS-CoV-2, particularly within the RBD of the Spike protein, which is a target for nAb, is a major concern associated with the use of first generation COVID-19 vaccines that typically target SARS-CoV-2 Spike protein.

[0078] To address these issues, a new booster vaccine was developed, c-srRNA-G5006d, which encodes a fusion protein comprising the CD5 signal peptide, Spike-RBD of SARS-CoV-2, nucleoprotein of SARS-CoV-2, nucleoprotein of MERS-CoV, and Spike-RBD of MERS-CoV (Example 12 and **FIG. 22**). The amino acid sequence of the pancoronavirus antigen (G5006d) is set forth as SEQ ID NO:27, and the nucleotide sequence of its open reading frame is set forth as SEQ ID NO:26. The order of each sequence segment (RBD of SARS-CoV-2; a nucleoprotein of SARS-CoV-2; a nucleoprotein of MERS-CoV; RBD of MERS-CoV) of the fusion protein can be

altered, and the amino acid sequences of each segment do not have to be 100% identical to the exemplary sequences provided herein.

[0079] The c-srRNA-G5006d vaccine is intended to be used as a booster vaccine, after a primary vaccine series (1st vaccination or 1st and 2nd vaccinations) targeted to the Spike antigen or fragment thereof (RBD) has been received. However, the c-srRNA-G5006d vaccine could also be used as part of a primary vaccine series.

[0080] The c-srRNA-G5006d vaccine boosts nAb levels and provides cellular immunity against betacoronaviruses that infect humans. Cellular immunity is important for providing long-lasting protection from severe illness, hospitalization, and death.

[0081] As described in Example 10, a c-srRNA vaccine encoding Spike-RBD can increase the level of antibodies or nAb against Spike-RBD, when it was used as a booster vaccine, following administration of a vaccine that can prime or induce humoral immunity.

[0082] The c-srRNA-G5006d encodes both Spike-RBD protein of SARS-CoV-2 and Spike-RBD protein of MERS-CoV. Therefore, c-srRNA-G5006d can be used as a booster vaccine for both SARS-CoV-2 and MERS-CoV.

[0083] Spike proteins of SARS-CoV-2 and SARS-CoV are similar (about 76% identity) (Grifoni et al., 2020). Therefore, c-srRNA-G5006d is effective as a booster for SARS-CoV-2, SARS-CoV, and their variants. On the other hand, Spike proteins of SARS-CoV-2 and MERS-CoV are different (about 35% identity) (Grifoni et al., 2020). However, c-srRNA-G5006d also encodes a Spike-RBD of MERS-CoV. Therefore, c-srRNA-G5006d is effective as a booster for MERS-CoV and its variants. Taken together, c-srRNA-G5006d is effective as a booster for SARS-CoV-2, SARS-CoV, MERS-CoV, and their variants.

[0084] The c-srRNA-G5006d also encodes nucleoproteins of SARS-CoV-2 and MERS-CoV. Therefore, c-srRNA-G5006d is able to induce strong cellular immunity against SARS-CoV-2 and MERS-CoV. Nucleoproteins of SARS-CoV-2 and SARS-CoV are very similar to each other (about 90% identity) (Grifoni et al., 2020). Therefore, c-srRNA-G5006d provides strong cellular immunity against SARS-CoV-2, SARS-CoV, and their variants. In contrast, nucleoproteins of SARS-CoV-2 and MERS-CoV are different (about 48% identity) (Grifoni et al., 2020). However, c-srRNA-G5006d also encodes a nucleoprotein of MERS-CoV. Therefore,

c-srRNA-G5006d is contemplated to provide strong cellular immunity against MERS-CoV and its variants. Taken together, c-srRNA-G5006dinduces a potent immune response against SARS-CoV-2, SARS-CoV, MERS-CoV, and their variants.

[0085] As described in Examples 9 and 10, c-srRNA vaccine has a remarkable mode of action. That is, the encoded antigens do not appear to directly stimulate B cells, and thus, consideration of three-dimensional structure of the encoded antigens is not required. This differs from traditional vaccine that are designed to directly stimulate the B cells to produce antibodies against conformational epitopes (three-dimensional structures of antigens). This is why it is appropriate to use a fusion protein for a c-srRNA vaccine, whereas use of a fusion protein for a traditional subunit vaccine is complicated by the fact that the natural three-dimensional structure of each antigen may be disrupted when expressed as a fusion protein. The c-srRNA booster vaccine stimulates antibody production through the activation of CD4+ helper T cells, and thus, it relies on short peptide epitopes (~15mer). Therefore, it is possible to simply put together two or more different antigens into a single fusion protein for an antigen encoded by a c-srRNA vaccine, while this mechanism may be problematic for design of a subunit vaccine.

[0086] The fact that c-srRNA relies on short peptide epitopes for induction of cellular and humoral immune responses also provides advantages for more broadly reactive vaccines that elicit protection against variant pathogens. Many T cell epitopes are present in a single protein, and thus, it is less likely that any single mutation will cause the loss of immunogenicity. On the other hand, traditional subunit vaccines rely on the three-dimensional structure of a protein antigen, and thus, even a single mutation may alter the conformation of the protein, which may lead to the loss of immunogenicity.

[0087] As shown in **FIG. 23A-23B**, c-srRNA-G5006d can stimulate cellular immunity against all proteins encoded by this vaccine: Spike-RBD of SARS-CoV-2, Nucleoprotein of SARS-CoV-2, Nucleoprotein of MERS-CoV, and Spike-RBD of MERS-CoV.

Pan-influenza booster vaccine

[0088] As determined during development of the present disclosure (see, e.g., Example 6), a fusion protein comprising nucleoproteins from representative Influenza A and Influenza B strains was able to induce a strong, antigen-specific cellular immune response when the fusion protein was expressed from an intradermally-injected, temperature-controllable, self-

replicating RNA. Protection is generally considered to be mainly mediated by neutralizing antibodies against hemagglutinin (HA), one of the surface proteins of influenza viruses. Therefore, FDA-approved influenza vaccines include HA as an antigen, alone or in combination with other influenza antigens. Since a c-srRNA-based booster vaccine requires only CD4+ T cell epitopes on the HA protein to enhance Ab production, the three-dimensional structure of the HA protein does not need to be considered. It is known that only some parts of the HA protein of the H1N1 influenza virus can function as CD4+ T cell epitopes (Knowlden et al., Pathogens. 8(4):220, 2019). B cell epitopes and CD4+ T cell epitopes in both influenza A and influenza B have been identified (Terajima et al. Virol J, 10:244, 2013). Sequences of HA proteins of representative H1N1 influenza viruses were aligned (Darricarrère et al., J Virol, 92(22):e01349-18, 2018) and regions with well-conserved sequences were identified. Based on these considerations, an HA protein fragment (residues 316-456) of Influenza A virus (A/New Caledonia/20/1999(H1N1)) [GenBank Accession No. EU103824] and an HA protein fragment (residues 332-474) of Influenza B virus (B/Florida/4/2006) [GenBank Accession No. CY033876] were selected. The nucleoproteins from Influenza A and Influenza B, which are already described in Example 6 and denoted as the G5010 antigen were also included.

[0089] **FIG. 24** shows the design of pan-influenza booster vaccine. The c-srRNA-G5012 encodes a fusion protein (G5012) comprising the signal peptide of CD5 (residues 1-24), a part of the hemagglutinin (HA) of the Influenza A, nucleoprotein of Influenza A, nucleoprotein of Influenza B, and a part of the hemagglutinin (HA) of the Influenza B. The amino acid sequence of the pan-influenza virus antigen (G5012) is set forth as SEQ ID NO:29, and the nucleotide sequence of its open reading frame is set forth as SEQ ID NO:28. The order of each sequence segment (a part of HA of Influenza A; a nucleoprotein Influenza A; a nucleoprotein of Influenza B; a part of HA of Influenza B) of the fusion protein can be altered, and the amino acid sequences of each segment do not have to be 100% identical to the exemplary sequences provided herein.

[0090] This c-srRNA-G5012 Influenza vaccine boosts nAb levels through the enhancement of HA-specific CD4+ helper T cells. It also provides cellular immunity against essentially all Influenza viruses through the evolutionary conserved nucleoproteins. The cellular

immunity is known to provide a long-lasting protection from severe illness, hospitalization, and death.

Chitosan-enhancement of gene expression in vivo

[0091] An RNase inhibitor (a protein purified from human placenta) slightly enhances the immunogenicity against an antigen encoded on c-srRNA, most likely by enhancing expression of the antigen from the c-srRNA in vivo when intradermally injected into mice (see e.g., FIG. 25C of WO 2021/138447 A1). The RNase inhibitor may protect c-srRNA from RNase-mediated degradation in vivo. However, it is desirable to find an alternative agent that can enhance expression of a gene of interest (GOI) in vivo for therapeutics purposes, as it is difficult to use a protein-based RNase inhibitor as an excipient in injectable products.

[0092] A low molecular weight chitosan (molecular weight ~ 6 kDa) was shown to inhibit the activity of RNase with the inhibition constants in the range of 30–220 nM (Yakovlev et al., Biochem Biophys Res Commun, 357(3):584-8, 2007). Although this has been shown only in vitro and also for artificially made poly nucleotides such as Poly(A)/Poly(U), whether chitosan oligosaccharides can enhance the expression of GOI from c-srRNA needed to be tested in vivo by intradermally injecting the c-srRNA in mice. As shown in Example 14, two different chitosan oligomers were tested: chitosan oligomer (molecular weight ≤ 5 kDa, ≥75% deacetylated: Heppe Medical Chitosan GmbH: Product No. 44009), and chitosan oligosaccharide lactate (molecular weight about 5 kDa, > 90% deacetylated: Sigma-Aldrich: Product No. 523682). Surprisingly, even a very low level of chitosan oligomers, as low as 0.001 μg/mL (about 0.2 nM: about 1/100 of the inhibition constant discovered by Yakovlev et al., supra, 2007) was found to be able to enhance the expression of luciferase encoded on c-srRNA by ~10-fold (**FIG. 25**). Similar enhancement of the GOI expression was achieved by chitosan oligomers for up to 0.5 μg/mL and by chitosan oligosaccharide lactate at 0.1 μg/mL.

[0093] Chitosan has been used as a nucleotide (DNA and RNA) delivery vector, as it can form complexes or nanoparticles (reviewed in Buschmann et al., Adv Drug Deliv Rev, 65(9):1234-70, 2013; and Cao et al., Drugs, 17:381, 2019). However, it is worth noting that the enhancement of the GOI expression by chitosan oligomers is unlikely to be mediated by the nanoparticle or the complex formation of c-srRNA and chitosan oligomers. First, such a low concentration of chitosan oligomers does not allow the complex formation with RNA. Second,

chitosan oligomers are added to c-srRNA immediately before the intradermal injection, and thus, there is not sufficient time to form the complex.

[0094] As the chitosan oligomers enhance expression of the GOI in vivo at much lower concentrations compared to the effective concentration as an RNase inhibitor in vitro (Yakovlev et al., supra, 2007), it is conceivable that this enhanced GOI expression by chitosan oligomers may not be mediated by its RNase inhibition mechanism. For example, chitosan oligomers may facilitate the incorporation of c-srRNA into cells, and thereby may enhance the expression of GOI from c-srRNA. Nonetheless, this surprising discovery should provide an effective means to enhance the in vivo therapeutic expression of GOI encoded on c-srRNA.

ENUMERATED EMBODIMENTS

1. A composition for stimulating an immune response against a coronavirus in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':
(i) a nucleotide sequence encoding a mammalian signal peptide; and
(ii) a nucleotide sequence encoding a coronavirus nucleocapsid protein.

2. The composition of embodiment 1, wherein the coronavirus is a betacoronavirus, optionally wherein the betacoronavirus is a human betacoronavirus.

3. The composition of embodiment 2, wherein the betacoronavirus comprises a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), a middle east respiratory syndrome-related coronavirus (MERS-CoV), or a combination thereof.

4. The composition of embodiment 3, wherein the betacoronavirus comprises a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2).

5. The composition of embodiment 4, wherein the coronavirus nucleocapsid protein comprises a first nucleocapsid protein and a second nucleocapsid protein, wherein the first nucleocapsid protein is a SARS-CoV-2 nucleocapsid protein of a first variant from a first clade, and the second nucleocapsid protein is a SARS-CoV-2 nucleocapsid protein of a second variant from a second clade, and wherein the first clade and the second clade are different clades as defined by one or more of the World Health Organization, Pango, GISAID, and Nextstrain.

6. A composition for stimulating an immune response against a coronavirus in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':
(i) a nucleotide sequence encoding a mammalian signal peptide; and
(ii) a nucleotide sequence encoding two or more coronavirus nucleocapsid proteins.

7. The composition of embodiment 6, wherein the coronavirus is a betacoronavirus, optionally wherein the betacoronavirus is a human betacoronavirus.

8. The composition of embodiment 7, wherein the betacoronavirus comprises a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), a middle east respiratory syndrome-related coronavirus (MERS-CoV), or a combination thereof.

9. The composition of embodiment 8, wherein the betacoronavirus comprises a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2).

10. The composition of embodiment 9, wherein the two or more coronavirus nucleocapsid proteins comprise a SARS-CoV-2 nucleocapsid protein and a MERS nucleocapsid protein.

11. The composition of embodiment 9, wherein the two or more coronavirus nucleocapsid proteins comprise a SARS-CoV-2 nucleocapsid protein, a SARS-CoV-1 nucleocapsid protein, and a MERS nucleocapsid protein.

12. The composition of any one of embodiments 6-11, wherein the two or more coronavirus nucleocapsid proteins are separated by a linker of from one to ten residues in length.

13. The composition of any one of embodiments 1-12, wherein the mammalian signal peptide is a signal peptide of a surface protein expressed in mammalian antigen presenting cells.

14. The composition of embodiment 13, wherein the mammalian signal peptide is a CD5 signal peptide and the amino acid sequence of the CD5 signal peptide comprises SEQ ID NO:8, or the amino acid sequence at least 90% or 95% identical to SEQ ID NO:8.

15. The composition of any one of embodiments 1-14, wherein the amino acid sequence of the nucleocapsid protein comprises residues 2-419 of SEQ ID NO:5, or the amino

acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to residues 2-419 of SEQ ID NO:5.

16. The composition of any one of embodiments 1-14, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:6, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:6.

17. The composition of any one of embodiments 6-14, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:7, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:7.

18. The composition of embodiment 16, wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO:2.

19. The composition of embodiment 17, wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO:3 or SEQ ID NO:4.

20. The composition of any one of embodiments 1-14, wherein the amino acid sequence of the fusion protein comprises residues 2-413 of SEQ ID NO:9, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to residues 2-413 of SEQ ID NO:9.

21. The composition of any one of embodiments 1-14, wherein the amino acid sequence of the fusion protein comprises residues 2-422 of SEQ ID NO:10, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to residues 2-422 of SEQ ID NO:10.

22. The composition of any one of embodiments 1-21, wherein the composition does not comprise liposomes or lipid nanoparticles.

23. The composition of any one of embodiments 1-22, wherein the mRNA is a self-replicating mRNA.

24. The composition of embodiment 23, wherein the self-replicating RNA comprises an Alphavirus replicon lacking a viral structural protein coding region.

25. The composition of embodiment 24, wherein the Alphavirus is selected from the group consisting of a Venezuelan equine encephalitis virus, a Sindbis virus, and a Semliki Forrest virus.

26. The composition of embodiment 25, wherein the Alphavirus is a Venezuelan equine encephalitis virus.

27. The composition of any one of embodiments 23-26, wherein the Alphavirus replicon comprises a nonstructural protein coding region with an insertion of 12-18 nucleotides resulting in expression of a nonstructural Protein 2 (nsP2) comprising from 4 to 6 additional amino acids between beta sheet 4 and beta sheet 6 of the nsP2.

28. The composition of any one of embodiments 1-27, wherein the self-replicating mRNA is a temperature-sensitive agent (ts-agent) that is capable of expressing the fusion at a permissive temperature but not at a non-permissive temperature.

29. The composition of embodiment 28, wherein the permissive temperature is from 31°C to 35°C and the non-permissive temperature is at least 37°C ± 0.5°C.

30. A method for stimulating an immune response against a coronavirus in a mammalian subject, comprising administering the composition of any one of embodiments 1-29 to a mammalian subject so as to stimulate an immune response against the coronavirus nucleocapsid protein in the mammalian subject

31. The method of embodiment 30, wherein the composition is administered intradermally.

32. The method of embodiment 30 or embodiment 31, wherein the immune response comprises a coronavirus-reactive cellular immune response.

33. The method of embodiment 32, wherein the immune response further comprises a coronavirus-reactive humoral immune response.

34. The method of any one of embodiments 30-33, wherein the mammalian subject is a human subject.

35. A kit comprising:
the composition of any one of embodiments 1-29 or any one of embodiments 37-62; and
a device for intradermal delivery of the composition to a mammalian subject.

36. The kit of embodiment 35, wherein the device comprises a syringe and a needle.

37. A composition for stimulating an immune response against two or more viruses in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':
(i) a nucleotide sequence encoding a mammalian signal peptide; and
(ii) a nucleotide sequence encoding a first nucleocapsid protein of a first virus and a second nucleocapsid protein of a second virus.

38. The composition of embodiment 37, wherein the first and second viruses are capable of causing disease upon infection of a human subject.

39. The composition of embodiment 38, wherein the first and second viruses are different variants, subtypes or lineages of the same species.

40. The composition of embodiment 38, wherein the first and second viruses are different species of the same genus.

41. The composition of embodiment 40, wherein the first and second viruses are both members of the betacoronavirus genus.

42. The composition of embodiment 41, wherein the first and second viruses comprise a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) and a middle east respiratory syndrome-related coronavirus (MERS-CoV).

43. The composition of embodiment 38, wherein the first and second viruses are members of different families, orders, classes, or phyla of the same kingdom.

44. The composition of embodiment 43, wherein the first and second viruses are both members of the orthomyxoviridae family.

45. The composition of embodiment 44, wherein the first and second viruses comprise an influenza A virus and an influenza B virus.

46. The composition of embodiment 45, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:16, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:16.

47. The composition of embodiment 38, wherein the first and second viruses are both members of the orthornavirae kingdom, optionally wherein the first and second viruses comprise: (a) a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), or a middle east respiratory syndrome-related coronavirus (MERS-CoV); and (b) an influenza A virus or an influenza B virus.

48. The composition of embodiment 40, wherein the first and second viruses are both members of the ebolavirus genus, optionally wherein the first and second viruses are selected from the group consisting of Zaire ebolavirus, Sudan ebolavirus, Bundibugyo ebolavirus, and Taï Forest ebolavirus.

49. The composition of embodiment 48, wherein the nucleotide sequence further encodes a third nucleocapsid protein of a third virus and a fourth nucleocapsid protein of a fourth virus, and the first, second, third and fourth viruses are Zaire ebolavirus, Sudan ebolavirus, Bundibugyo ebolavirus, and Taï Forest ebolavirus.

50. The composition of embodiment 49, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:22, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:22.

51. The composition of embodiment 49, wherein the nucleotide sequence (ii) encodes a shared portion of the first nucleocapsid protein of the first virus for stimulating an immune response against all of the first, second, third and fourth viruses.

52. The composition of embodiment 51, wherein the nucleotide sequence (ii) encodes an individual portion of each of the first, second, third and fourth nucleocapsid proteins for stimulating an immune response against all of the first, second, third and fourth viruses.

53. The composition of embodiment 52, wherein the nucleotide sequence (ii) encodes a fragment of the individual portion of the second nucleocapsid protein of the second virus for stimulating an immune response against the second and third viruses.

54. The composition of embodiment 37, wherein the nucleotide sequence (ii) encodes a shared portion of the first nucleocapsid protein of the first virus for stimulating an immune response against both the first and second viruses.

55. The composition of embodiment 54, wherein the nucleotide sequence (ii) encodes an individual portion of each of the first and second nucleocapsid proteins for stimulating an immune response against both the first and second viruses.

56. The composition of any one of embodiments 37-48, wherein the nucleotide sequence of (ii) further encodes at least one further nucleocapsid protein of at least one further virus, and wherein the at least one further virus is different from the first and second viruses.

57. The composition of any one of embodiments 37-56, wherein the first and second, or the first, second, and further nucleocapsid proteins are separated by a linker of from one to ten residues in length.

58. The composition of any one of embodiments 37-57, wherein the mammalian signal peptide is a signal peptide of a surface protein expressed in mammalian antigen presenting cells.

59. The composition of any one of embodiments 37-58, wherein the mRNA is a self-replicating mRNA.

60. The composition of embodiment 59, wherein the self-replicating mRNA is a temperature-sensitive agent (ts-agent) that is capable of expressing the fusion protein a permissive temperature but not at a non-permissive temperature.

61. The composition of embodiment 60, wherein the permissive temperature is from 31°C to 35°C and the non-permissive temperature is at least 37°C ± 0.5°C.

62. The composition of any one of embodiments 1-29 or any one of embodiments 37-61, wherein the composition further comprises chitosan.

63. A method for stimulating an immune response against two or more viruses in a mammalian subject, comprising administering the composition of any one of embodiments 37-62 to a mammalian subject to stimulate an immune response against the nucleocapsid proteins of the two or more viruses in the mammalian subject

64. The method of embodiment 63, wherein the composition is administered intradermally.

65. The method of embodiment 63 or embodiment 64, wherein the immune response comprises a cellular immune response reactive with the two or more viruses.

66. The method of embodiment 65, wherein the cellular immune response comprises a nucleocapsid protein-specific helper T lymphocyte (Th) response comprising nucleocapsid protein-specific cytokine secretion.

67. The method of embodiment 66, wherein nucleocapsid protein-specific cytokine secretion comprises secretion of one or both of interferon-gamma and interleukin-4.

68. The method of embodiment 65, wherein the cellular immune response comprises a nucleocapsid protein-specific cytotoxic T lymphocyte (CTL) response.

69. The method of any one of embodiments 65-68, wherein the immune response further comprises a humoral immune response reactive with the two or more viruses.

70. The method of any one of embodiments 63-69, wherein the mammalian subject is a human subject.

71. A composition for stimulating an immune response against a virus in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':

(i) a nucleotide sequence encoding a mammalian signal peptide;

(ii) a nucleotide sequence encoding a first viral antigen or fragment thereof of a first virus; and

(iii) a nucleotide sequence encoding a second viral antigen or fragment thereof of the first virus or a second virus,

wherein the first viral antigen is a nucleocapsid protein and the second viral antigen is a surface protein, or the first viral antigen is a surface protein and the second viral antigen is a nucleocapsid protein.

72. A composition for stimulating an immune response against two or more viruses in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':

(i) a nucleotide sequence encoding a mammalian signal peptide;

(ii) a nucleotide sequence encoding a first viral antigen or fragment thereof of a first virus;

(iii) a nucleotide sequence encoding a second viral antigen or fragment thereof of the first virus;

(iv) a nucleotide sequence encoding a third viral antigen or fragment thereof of a second virus;

(iii) a nucleotide sequence encoding a fourth viral antigen or fragment thereof of the second virus,

wherein the first viral antigen is a first nucleocapsid protein and the second viral antigen is a first surface protein, or the first viral antigen is a first surface protein and the second viral antigen is a first nucleocapsid protein, and

wherein the third viral antigen is a second nucleocapsid protein and the fourth viral antigen is a second surface protein, or the third viral antigen is a second surface protein and the fourth viral antigen is a second nucleocapsid protein.

73. The composition of embodiment 71 or embodiment 72, wherein the mRNA is a self-replicating mRNA.

74. The composition of embodiment 73, wherein the self-replicating RNA comprises an Alphavirus replicon lacking a viral structural protein coding region.

75. The composition of embodiment 74, wherein the Alphavirus is selected from the group consisting of a Venezuelan equine encephalitis virus, a Sindbis virus, and a Semliki Forrest virus.

76. The composition of embodiment 74, wherein the Alphavirus is a Venezuelan equine encephalitis virus.

77. The composition of any one of embodiments 73-76, wherein the self-replicating mRNA is a temperature-sensitive agent (ts-agent) that is capable of expressing the fusion protein at a permissive temperature but not at a non-permissive temperature.

78. The composition of embodiment 77, wherein the permissive temperature is from 31°C to 35°C, and the non-permissive temperature is at least 37°C ± 0.5°C.

79. The composition of any one of embodiments 74-78, wherein the Alphavirus replicon comprises a nonstructural protein coding region with an insertion of 12-18 nucleotides resulting in expression of a nonstructural Protein 2 (nsP2) comprising from 4 to 6 additional amino acids between beta sheet 4 and beta sheet 6 of the nsP2.

80. The composition of any one of embodiments 71-79, wherein the first virus and/or the second virus is a coronavirus, optionally wherein the coronavirus is a betacoronavirus, optionally wherein the betacoronavirus is a human betacoronavirus.

81. The composition of embodiment 80, wherein the first and/or the second virus is a betacoronavirus independently selected from the group consisting of a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), and a middle east respiratory syndrome-related coronavirus (MERS-CoV).

82. The composition of embodiment 80, wherein the first virus is SARS-CoV-2 and the second virus is MERS-CoV.

83. The composition of any one of embodiments 80-82, wherein the surface protein, the first surface protein and/or the second surface protein each comprise a receptor-binding domain (RBD) of a coronavirus Spike protein.

84. The composition of embodiment 83, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:27, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:27.

85. The composition of any one of embodiments 71-79, wherein the first virus and/or the second virus is a member of the orthomyxoviridae family.

86. The composition of embodiment 85, wherein the first and/or the second virus is independently selected from the group consisting of an influenza A virus (IAV) and an influenza B virus (IBV).

87. The composition of embodiment 86, wherein the first virus is IAV and the second virus is IBV.

88. The composition of any one of embodiments 85-87, wherein the surface protein, the first surface protein and/or the second surface protein each comprise a portion of an influenza hemagglutinin.

89. The composition of embodiment 88, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:29, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:29.

90. The composition of any one of embodiments 71-89, wherein the composition further comprises chitosan.

91. A kit comprising:
(i) the composition of any one of embodiments 71- 90; and
(ii) a device for intradermal delivery of the composition to a mammalian subject.

92. The kit of embodiment 91, wherein the device comprises a syringe and a needle.

93. The kit of embodiment 91 or embodiment 92, further comprising instructions for use of the device to administer the composition to a mammalian subject to stimulate an immune response against one or more of the first viral antigen, the second viral antigen, the third viral antigen, and the fourth viral antigen.

94. A method of stimulating an immune response in a mammalian subject, comprising administering the composition of any one of embodiments 71-90 to a mammalian subject to stimulate an immune response against one or more of the first viral antigen, the second viral antigen, the third viral antigen, and the fourth viral antigen in the mammalian subject.

95. The method of embodiment 94, wherein the composition is administered intradermally.

96. The method of embodiment 95, wherein the immune response comprises a cellular immune response reactive against one or more of the first viral antigen, the second viral antigen, the third viral antigen, and the fourth viral antigen.

97. The method of embodiment 96, wherein the immune response further comprises a humoral immune response reactive against one or more of the first viral antigen, the second viral antigen, the third viral antigen, and the fourth viral antigen.

98. The method of any one of embodiments 94-97, wherein the mammalian subject is a human subject.

99. A method for active booster immunization against at least one virus, comprising intradermally administering the composition of any one of embodiments 1-29, any one of embodiments 37-62, or any one of embodiments 71-90 to a mammalian subject in need thereof to stimulate a secondary immune response against the virus, wherein the mammalian subject had already undergone a primary immunization regimen against the virus.

100. The method of embodiment 99, wherein the primary immunization regimen comprises administration of at least one dose of a different vaccine against the virus.

101. The method of embodiment 100, wherein the different vaccine comprises a protein antigen of the at least one virus, optionally wherein the protein antigen is a recombinant protein or fragment thereof, or an inactivated virus.

102. A method for active booster immunization against at least one virus, comprising:

(i) intradermally administering the composition of any one of embodiments 1-29, any one of embodiments 37-62, or any one of embodiments 71-90 to a mammalian subject in need thereof to stimulate a primary immune response against the virus; and

(ii) administering at least one dose of a different vaccine against the virus to the mammalian subject to stimulate a secondary immune response against the virus.

103. The method of embodiment 102, wherein the different vaccine comprises a protein antigen of the at least one virus, optionally wherein the protein antigen is a recombinant protein or fragment thereof, or an inactivated virus.

104. A method for active primary immunization against at least one virus, comprising:

(i) intradermally administering the composition of any one of embodiments 1-29, any one of embodiments 37-62, or any one of embodiments 71-90 to a mammalian subject in need thereof to stimulate a primary immune response against the virus; wherein the mammalian subject had not undergone a primary immunization regimen against the virus.

105. The method of embodiment 104, further comprising:

(ii) administering at least one dose of a different vaccine against the virus to the mammalian subject to stimulate a secondary immune response against the virus.

106. The method of embodiment 105, wherein the different vaccine comprises a protein antigen of the at least one virus, optionally wherein the protein antigen is a recombinant protein or fragment thereof, or an inactivated virus.

107. The method of any one of embodiments 94-106, wherein the mammalian subject is a human subject.

108. An expression vector comprising the mRNA of any of the preceding claims in operable combination with a promoter.

109. The expression vector of embodiment 108, wherein the promoter is a T7 promoter or a SP6 promoter.

110. The expression vector of embodiment 108, wherein the vector is a plasmid.

111. The expression vector of any one of embodiments 108-110, further comprising a selectable marker.

EXAMPLES

[0095] Abbreviations: Ab (antibody); APC (antigen presenting cell); CoV (coronavirus); c-srRNA (temperature-controllable, self-replicating RNA); CTL (cytotoxic T lymphocyte); FluA or IAV (influenza A virus); FluB or IBV (influenza B virus); IL-4 (interleukin-4); INF-γ (interferon gamma); GOI (gene of interest); HA (hemagglutinin); MERS (middle east respiratory syndrome-related); nAb (neutralizing antibody); N or NP (nucleocapsid or nucleoprotein); nsP (non-structural protein); ORF (open reading frame); PBO (placebo); RBD (receptor-binding domain); S (spike); PRNT (plaque reduction neutralization test); SARS (severe acute respiratory syndrome); SFC (spot-forming cells); SFU (spot-forming units); srRNAts (temperature-sensitive, self-replicating RNA); Th (helper T lymphocyte); and Tx (treatment). The terms c-srRNA and srRNAts are used interchangeably throughout the disclosure, with srRNA1ts2 (described in WO 2021/138447 A1) being an exemplary embodiment.

Example 1. Cellular immunity induced by srRNA1ts2-G5004

[0096] This example describes the finding that SARS-CoV-2 nucleoprotein alone (G5004 antigen, without a signal peptide) does not induce a potent cellular immune response when the protein is expressed from intradermally-injected, temperature-controllable, self-replicating RNA.

Materials and Methods.

[0097] CD-1 outbred female mice.

[0098] srRNA1ts2-G5004 mRNA was produced by in vitro transcription of a temperature-controllable, self-replicating RNA vector (srRNA1ts2 as described in PCT/US2020/067506) encoding the G5004 antigen (**FIG. 2**).

[0099] A pool of 102 peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein (UniProt: P0DTC9) of SARS-CoV-2 [JPT peptide Product Code: PM-WCPV-NCAP].

[0100] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0101] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

Results

[0102] Recently, it has been shown that vaccination with nucleoprotein (N) alone elicits cellular immunity and spike-independent SARS-CoV-2 protective immunity in mice and hamsters (Machett et al., bioRxiv. 2021.04.26.441518.2021). Vaccination involved intravenous administration of a human adenovirus serotype 5 (Ad5) vector expressing the N sequence (Ad5-N) derived from USA-WA1/2021 strain.

[0103] To test whether nucleoprotein (N) alone (without a signal peptide) can induce cellular immunity, ELISpot assays were performed 14 days after vaccinating CD-1 outbred mice by a single intradermal injection of either 5 μg or 25 μg of an srRNA1ts2-G5004 (**FIG. 2**) or a placebo (PBO: buffer only). Only weak induction of interferon-gamma (INF-γ)-secreting T cells (**FIG. 4A**) and IL-4-secreting T cells (**FIG. 4B**) was observed. Interestingly, the INF-γ response was not observed to be dose-dependent (5 μg vs. 25 μg).

[0104] It was concluded that the nucleoprotein (N) alone did not induce a potent cellular immune response when expressed from the intradermally-injected, temperature-controllable, self-replicating RNA.

Example 2. Cellular immunity induced by srRNA1ts2-G5005

[0105] This example describes the finding that the addition of a CD5-signal peptide to SARS-CoV-2 nucleoprotein induces a potent cellular immune response in CD-1 mice when expressed from intradermally-injected, temperature-controllable, self-replicating RNA.

Materials and Methods

[0106] CD-1 outbred female mice.

[0107] srRNA1ts2-G5005 mRNA was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 as disclosed in PCT/US2020/067506]) encoding the G5005 antigen (**FIG. 2**).

[0108] A pool of 102 peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein (UniProt: P0DTC9) of SARS-CoV-2 [JPT peptide Product Code: PM-WCPV-NCAP].

[0109] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0110] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

Results

[0111] The wild-type nucleoprotein does not contain a signal peptide or a transmembrane domain, and therefore is not expected to be directed to the mammalian host cell's secretory pathway. The inventor reasoned that the lack of a signal peptide may be why the wild-type nucleoprotein (expressed from srRNA1ts2-G5004 of Example 1) did not induce a potent cellular immune response. With this in mind, the coding region of the signal peptide sequence from the human CD5 gene was added to the nucleoprotein coding region in place of the start codon (ATG) of the nucleoprotein in srRNA1ts2-G5005 (**FIG. 2**). The amino acid sequence of the CD5 signal peptide is MPMGSLQPLATLYLLGMLVASCLG (set forth as SEQ ID NO:8).

[0112] Cellular immunity was assessed by ELISpot assays 14 days after vaccinating CD-1 outbred mice by a single intradermal injection of either 5 μg or 25 μg of an srRNA1ts2-G5005 (**FIG. 2**) or a placebo (PBO: buffer only).

[0113] As shown in **FIG. 5A**, antigen-specific, INF-γ-secreting T cells were strongly induced in a dose-dependent manner (5 μg vs. 25 μg). By contrast, there was little to no induction of antigen-specific IL-4-secreting T cells (**FIG. 5B**). Th1 cells secrete INF-γ, while Th2 cells secrete IL-4. It is generally accepted that a Th1>Th2 immune response is a favorable feature of a vaccine.

[0114] In conclusion, addition of a signal peptide derived from human CD5 to the N-terminus of the nucleoprotein (N) resulted in induction of a strong antigen-specific cellular immune response when the protein is expressed from intradermally-injected, temperature-controllable, self-replicating RNA. The srRNA1ts2-G5005 vaccine also showed a favorable Th1-skewed (Th1>Th2) immune response.

Example 3. Cellular immunity induced by srRNA1ts2-G5005

[0115] This example describes the finding that the addition of a CD5-signal peptide to the SARS-CoV-2 nucleoprotein induces a potent cellular immune response in BALB/c mice when expressed from intradermally-injected, temperature-controllable, self-replicating RNA.

Materials and Methods

[0116] BALB/c female mice.

[0117] srRNA1ts2-G5005 mRNA was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 as described in PCT/US2020/067506]) encoding the G5005 antigen (**FIG. 2**).

[0118] A pool of 102 peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein (UniProt: P0DTC9) of SARS-CoV-2 [JPT peptide Product Code: PM-WCPV-NCAP].

[0119] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0120] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

Results

[0121] To test whether srRNA1ts2-G5005 can induce a strong cellular immune response in another mouse strain, an immunogenicity study was also conducted in BALB/c mice. Cellular immunity was assessed by ELISpot assays 30 days after vaccinating BALB/c mice by a single intradermal injection of either 5 μg or 25 μg of srRNA1ts2-G5005 (**FIG. 2**) or a placebo (PBO: buffer only).

[0122] As shown in **FIG. 6A**, antigen-specific, INF-γ-secreting T cells was strongly induced in a dose-dependent manner (5 μg vs. 25 μg). By contrast, antigen-specific, IL-4-secreting T cells were not induced (**FIG. 6B**). Therefore, a favorable Th1>Th2 cellular response was also observed in BALB/c mice.

[0123] In conclusion, addition of a signal peptide derived from human CD5 to the N-terminus of the nucleoprotein (N) significantly enhanced an antigen-specific cellular immune response when the protein is expressed from intradermally-injected, temperature-controllable, self-replicating RNA. As in CD-1 mice, the srRNA1ts2-G5005 vaccine showed a favorable Th1 skewed (Th1>Th2) immune response in BALB/c mice.

Example 4. Humoral immunity induced by srRNA1ts2-G5005

[0124] This example describes the finding that the SARS-CoV-2 nucleoprotein when linked to the human CD5-signal peptide induces a potent humoral immune response when the protein is expressed from intradermally-injected, temperature-controllable, self-replicating RNA.

Materials and Methods

[0125] BALB/c female mice.

[0126] srRNA1ts2-G5005 mRNA was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 as described in PCT/US20/67506) encoding the G5005 antigen (**FIG. 2**).

[0127] SARS-CoV-2 Nucleocapsid IgG ELISA kit (ENZO: ENZ-KIT193-0001).

Results

[0128] To test whether srRNA1ts2-G5005 can induce a humoral immunity, nucleoprotein-specific IgG levels in serum was measured by ELISA 30 days after vaccinating BALB/c mice by a single intradermal injection of either 5 μg or 25 μg of srRNA1ts2-G5005 (**FIG. 2**) or a placebo (PBO: buffer only). The IgG levels are represented by OD450 in the ELISA. The IgG levels were measured before (Day -1) and after (Day 30) vaccination (Day 0).

[0129] As shown in **FIG. 7**, nucleoprotein-specific serum IgG was strongly induced in a dose-dependent manner (5 μg vs. 25 μg).

[0130] In conclusion, addition of a signal peptide derived from human CD5 to the N-terminus of the nucleoprotein (N) induced an antigen-specific humoral immune response when the protein is expressed from intradermally-injected, temperature-controllable, self-replicating RNA.

Example 5. Cellular immunity induced by srRNA1ts2-G5006

[0131] This example describes the finding that a fusion protein comprising the SARS-CoV-2 nucleoprotein and the MERS-CoV nucleoprotein can induce strong cellular immunity against SARS-CoV-2 and MERS-CoV when the protein is expressed from intradermally-injected, temperature-controllable, self-replicating RNA.

Materials and Methods

[0132] BALB/c female mice.

[0133] srRNA1ts2-G5006 mRNA was produced by in vitro transcription of a temperature-controllable, self-replicating, RNA vector (srRNA1ts2 as described in PCT/US2020/067506) encoding the G5006 antigen (**FIG. 2C**).

[0134] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein (UniProt: P0DTC9) of SARS-CoV-2 [JPT peptide Product Code: PM-WCPV-NCAP].

[0135] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein of MERS-CoV.

[0136] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0137] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

Results

[0138] T-cell epitopes are present in short linear peptides, typically within the size range of 8-11 residues for MHC class I, and 10-30 residues for MHC class II. Unlike many B-cell epitopes, the 3-D conformation of T-cell epitopes is not critical to recognition by immune cell receptors. Therefore, the inventor reasoned that nucleoproteins from different betacoronavirus strains can be fused together in the absence of a lengthy linker (greater than 10 amino acids in length) for use as a vaccine antigen to elicit an immune response against different betacoronaviruses (e.g., SARS-CoV-1 and their variants, SARS-CoV-2 and their variants, and MERS-CoV and their variants).

[0139] To test this concept, a fusion protein comprising a human CD5-signal peptide, a SARS-CoV-2 nucleoprotein, and a MERS-CoV nucleoprotein was designed (see G5006 in **FIG. 2C**). Mice were vaccinated with srRNA1ts2-G5006 by intradermal injection, and antigen-specific cellular immune responses were measured by ELISpot assays. As expected, the srRNA1ts2-G5006 vaccine induced a strong INF-γ-secreting T cell response against both the

SARS-CoV-2 nucleoprotein (**FIG. 8**) and the MERS-CoV nucleoprotein. Additionally, the cellular immune response is expected to have a Th1>Th2 balance.

[0140] In conclusion, a fusion protein comprising nucleoproteins from different betacoronaviruses induced a strong, antigen-specific cellular immune response when the fusion protein is expressed from intradermally-injected, temperature-controllable, self-replicating RNA.

Example 6. Cellular immunity induced by srRNA1ts2-G5010 (pan-Influenza vaccine)

[0141] This example describes the assessment of the immune response induced by a fusion protein comprising an Influenza A virus (FluA) nucleoprotein and an Influenza B virus (FluB) nucleoprotein when the protein is expressed from an intradermally injected temperature-controllable self-replicating RNA.

Materials and Methods

[0142] BALB/c female mice.

[0143] srRNA1ts2-G5010 mRNA was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 [PCT/US20/67506]) encoding the G5010 antigen (**FIG. 9**). The amino acid sequence of the G5010 fusion protein is set forth as SEQ ID NO:16. The nucleic acid sequence encoding the G5010 fusion protein was codon-optimized for expression in human cells, and is set forth as SEQ ID NO:15.

[0144] A pool of 122 overlapping peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein (NP) of Influenza A (H2N2) (Swiss-Prot ID P21433) [JPT peptide Product Code: PM-INFA-NPH2N2]. The amino acid sequence of the H2N2 nucleoprotein is set forth as SEQ ID NO:17.

[0145] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0146] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

Results

[0147] Influenza A and B can infect humans and cause seasonal epidemics or pandemics (see, "Types of Influenza Viruses" from the CDC website www.cdc.gov/flu/about/viruses/types.htm). Compared to the hemagglutinin (HA) and

neuraminidase (NA) antigens that are routinely included in influenza vaccines, the nucleoprotein antigens are more conserved among different Influenza virus strains. For example, the amino acid sequences of nucleoproteins of representative Influenza A strains (H1N1, H3N2, H5N8, H7N7, H7N9, H9N2, H10N8) are very similar. Likewise, the amino acid sequences of nucleoproteins of representative Influenza B strains (Yamagata, Victoria) are very similar. In contrast, the amino acid sequences of nucleoproteins of Influenza A are significantly different from the amino acid sequences of nucleoproteins of Influenza B.

[0148] T-cell epitopes are present in short linear peptides, typically within the size range of 8-11 residues for MHC class I and 10-30 residues for MHC class II. Unlike B-cell epitopes, the conformational or 3D structure of T-cell epitopes is not critical to recognition by immune cell receptors. Therefore, one representative nucleoprotein from Influenza A is contemplated to include many T-cell epitopes shared by many Influenza A virus strains. Likewise, one representative nucleoprotein from Influenza B is contemplated to include many T-cell epitopes shared by many Influenza B virus strains. As such, the inventor reasoned that the nucleoproteins from different Influenza strains can be fused together in the absence of a lengthy linker (greater than 10 amino acids in length) for use as a vaccine antigen to elicit immune responses against different Influenza viruses (e.g., different strains of Influenza A, and different strains of Influenza B).

[0149] The amino acid sequences of nucleoproteins of representative Influenza A strains (H1N1, H3N2, H5N8, H7N7, H7N9, H9N2, and H10N8) were found to be similar to each other. The nucleoprotein of Influenza strain H5N8 was selected as it showed the fewest differences to the nucleoproteins of other strains (H1N1, H3N2, H7N7, H7N9, H9N2, and H10N8). The nucleoprotein of Influenza B strain (B/Florida/4/2006; GenBank CY033879.1) was selected as a representative Influenza B virus nucleoproteins. A fusion protein comprising a human CD5-signal peptide, one FluA nucleoprotein and one FluB nucleoprotein was designed (see, G5010 in **FIG. 9**), and the coding region of the fusion protein was cloned downstream of the subgenomic promoter of srRNA1ts2. mRNA was subsequently produced by in vitro transcription. The amino acid sequence of the FluA nucleoprotein is set forth as SEQ ID NO:13 (Influenza Type A, H5N8 subtype [A/breeder duck/Korea/Gochang1/2014], GenBank No. KJ413835.1, ProteinID No. AHL21420.1), and the amino acid sequence of the FluB nucleoprotein is set forth as SEQ ID

NO:14 (Influenza Type B [B/Florida/4/2006], GenBank No. CY033879.1, ProteinID No. ACF54251.1).

[0150] Mice were vaccinated with srRNA1ts2-G5010 by intradermal injection, and antigen-specific cellular immune responses were measured by ELISpot assays. In order to recall nucleoprotein-reactive T cell immunity, a pool of 122 overlapping peptides derived from a peptide scan of the Influenza A nucleoprotein sequence set forth as SEQ ID NO:17 were used to restimulate splenocytes harvested from mice 14 days post-vaccination. Even though, there were differences between the influenza A nucleoprotein sequence of G5010 and the influenza A nucleoprotein sequence of the peptide pool (**FIG. 10**), the srRNA1ts2-G5010 vaccine induced a strong INF-γ-secreting T cell response against the FluA nucleoprotein (**FIG. 11**). Importantly, there was little to no induction of IL-4-secreting T cells against the FluA nucleoprotein. These results indicate that the srRNA1ts2-G5010 vaccine induces a Th1 (INF-γ)-dominant response (Th1>Th2 balance), which is a favorable feature for a vaccine directed against a viral disease.

[0151] In conclusion, a fusion protein comprising nucleoproteins from representative Influenza A and Influenza B strains induced a strong, antigen-specific cellular immune response when the fusion protein was expressed from intradermally-injected, temperature-controllable, self-replicating RNA.

Example 7. Cellular immunity induced by srRNA1ts2-PanEbola (pan-Ebola vaccine)

[0152] This example describes the finding that a fusion protein comprising fragments of nucleoproteins from four species of Ebolavirus (Zaire ebolavirus, Sudan ebolavirus, Bundibugyo ebolavirus, Taï Forest ebolavirus) can induce strong cellular immunity against Ebolaviruses when the fusion protein is used as a vaccine antigen. This example uses a temperature-controllable self-replicating RNA as an expression vector.

Materials and Methods

[0153] BALB/c female mice.

[0154] srRNA1ts2-PanEbola mRNA was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 [PCT/US20/67506]) encoding a PanEbola antigen (**FIG. 12**).

[0155] A pool of 182 peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein (Swiss-Prot ID: B8XCN6) of Ebola virus - Taï Forest Ebolavirus [JPT peptide; PepMix Tai Forest Ebolavirus (NP); JPT Product Code: PM-TEBOV-NP].

[0156] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0157] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

Results

[0158] Ebolaviruses cause highly lethal hemorrhagic fever. Four species of Ebolavirus are known to cause disease in humans: Ebola virus (species Zaire ebolavirus), Sudan virus (species Sudan ebolavirus), Bundibugyo virus (species Bundibugyo ebolavirus), and Taï Forest virus (species Taï Forest ebolavirus, formerly Côte d'Ivoire ebolavirus).

[0159] Currently, only one licensed vaccine (rVSV-ZEBOV) is available for Ebolavirus. This vaccine is an attenuated recombinant vesicular stomatitis virus (VSV), which expresses the main glycoprotein (GP) from the Zaire ebolavirus. Although the vaccine can induce a neutralizing antibody against Ebolavirus, the protein sequence of the GP is highly divergent among the four species of Ebolavirus, which infect humans. As such, the rVSV-ZEBOV vaccine is only effective against the Zaire ebolavirus. It is desirable to have a pan-ebolavirus vaccine, which could provide protection against all four species of ebolaviruses.

[0160] Compared to GP, the nucleoprotein (NP) sequences are more conserved among the four species of ebolavirus. However, unlike the GP, the NP is not a surface protein, and thus, the antibody induced against NP is not a neutralizing antibody. Importantly, it has been shown that mice vaccinated against Zaire ebolavirus NP can be protected from the Zaire ebolavirus challenge, which is mediated by cellular immunity, not humoral immunity (Wilson and Hart, J Virol, 75:2660-2664, 2001). It has also been shown that protection is mediated by MHC class I-restricted CD8+ killer T cells (cytotoxic T lymphocytes), not by MHC class II-restricted CD4+ helper T cells (Wilson and Hart, supra, 2001).

[0161] Using a fusion protein of NPs of all four species of ebolavirus as a vaccine antigen provides was reasoned to provide protection against all four species of ebolavirus.

However, each NP is approximately 740 amino acids in length. Thus fusing four whole NPs together would result in a relatively large protein of approximately 3,000 amino acids. A smaller-sized antigen is desirable for many vaccine platforms.

[0162] The amino acid sequences of nucleoproteins of four ebolavirus species was compared using NCBI BlastP (Zaire ebolavirus NP (GenBank ID: AF272001), Sudan ebolavirus NP (GenBank ID: AF173836), Bundibugyo ebolavirus NP (GenBank ID: FJ217161), and Taï Forest ebolavirus NP (GenBank ID: FJ217162)). The sequences of the N-terminal half of NP (termed Region A) were found to be similar to each other (88%-92% identity), whereas the sequences of the C-terminus half of NP (termed Region B) were found to be diverse (42%-54%) (**Table 7-1**). Therefore, Zaire (A) was chosen as a representative of Zaire (A), Sudan (A), Bundibugyo (A), and Taï Forest (A). For Region B, the Bundibugyo (B) and Taï Forest (B) were found to be similar to each other (80% and 86% identity), except for the middle part (40% identity) (termed Region C). Therefore, Zaire (B), Sudan (B), Bundibugyo (B), and Taï Forest (C) were selected for inclusion in the Pan-Ebola vaccine. Before assembling the four nucleoproteins into a single fusion protein, an additional 8 amino acid sequence was added to both sides, so that possible T-cell epitopes at the end of the nucleoprotein fragments, would not be destroyed. A schematic of the fusion protein of the Pan-Ebola antigen is shown in **FIG. 12**, and includes NP fragments of Zaire (A), Zaire (B), Sudan (B), Bundibugyo (B), and Taï Forest (C), as well as the human CD5 signal peptide. A diagram showing percent identities of ebolavirus NP sequences is shown in **FIG. 13.** The amino acid sequence of the PanEbola antigen is set forth as SEQ ID NO:22, while the nucleic acid sequence encoding the PanEbola antigen is set forth as SEQ ID NO:23.

Table 7-1. Percent Identity Between Domains of Ebolavirus NP Sequences

Virus	Zaire (A)	Zaire (B)
Sudan	88%	42%
Bundibugyo	92%	53%
Tai Forest	92%	

[0163] The srRNA1ts2-PanEbola vaccine was produced by cloning the PanEbola fusion protein downstream of the subgenomic promoter of a srRNA1ts2. mRNA was produced by in

vitro transcription, and used to vaccinate BALB/c mice intradermally. Antigen-specific cellular immune responses were measured by ELISpot assays. In order to recall nucleoprotein-reactive T cell immunity, a pool of 182 peptides derived from a peptide scan of the nucleoprotein ((Swiss-Prot ID: B8XCN6) of Ebola virus - Taï Forest Ebolavirus)) were used to restimulate splenocytes harvested from mice 14 days post-vaccination. The srRNA1ts2-PanEbola vaccine induced a strong INF-γ-secreting T cell response against the Taï Forest nucleoprotein (**FIG. 14A**). This is striking in that only a small part (169 aa) of the Taï Forest nucleoprotein was included in the mRNA vaccine, whereas the peptide pool used for restimulation covered the entire Taï Forest nucleoprotein sequence. Importantly, there was little to no induction of IL-4-secreting T cells against the Taï Forest nucleoprotein (**FIG. 14B**). These results indicate that the srRNA1ts2-PanEbola vaccine induces a Th1 (INF-γ)-dominant response (Th1>Th2 balance), which is a favorable feature for a vaccine directed against a viral disease.

[0164] In conclusion, a fused protein comprising nucleoproteins fragments from four species of Ebolavirus induced a strong, antigen-specific cellular immune response when the fusion protein was expressed from intradermally-injected, temperature-controllable, self-replicating RNA. The example demonstrates that the size of a fusion protein to be used as a Pan-Ebola vaccine can be reduced by removing the more well-conserved portions of one or more of the nucleoproteins comprising the vaccine. The PanEbola antigen is also suitable for use in other vaccine platforms (e.g., adenovirus, adeno-associated virus, recombinant protein, etc.).

Example 8. Cellular immunity induced by srRNA1ts2-G5003o (omicron vaccine)

[0165] This example describes the finding that intradermal delivery of c-srRNA encoding the RBD of SARS-CoV-2 (omicron strain B.1.1.529) can induce strong cellular immunity in mice.

Materials and Methods

[0166] C57BL/6 female mice.

[0167] An srRNA1ts2-G5003o (mRNA), which was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 [WO 2021/138447 A1]) encoding the G5003o antigen (**FIG. 15**).

[0168] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through RBD of SARS-CoV-2 Omicron variant (S-RBD B.1.1.529) [JPT Peptides: PM-SARS2-RBDMUT08-1]

[0169] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0170] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

Results

[0171] In this example, c-srRNA encoding the RBD of SARS-CoV-2 omicron variant (G5003o) was generated (**FIG. 15**). The RNA was intradermally administered to C57BL/6 mice and 14 days later the splenocytes were collected to examine the cellular immunity against SARS-CoV-2 RBD (omicron variant). Induction of INF-γ-secreting T cells was specifically observed in c-srRNA-G5003o recipients (**FIG. 16A**), whereas induction of IL-4-secreting T cells was not observed in c-srRNA-G5003o recipients (**FIG. 16B**).

Conclusion

[0172] We have demonstrated by using an omicron variant-specific RBD as an antigen, that an omicron variant-specific cellular immune response can be induced when the protein is expressed from the intradermally injected temperature-controllable self-replicating RNA. A favorable Th1 (INF-γ) > Th2 (IL-4) response was also observed.

Example 9. Efficacy of c-srRNA Prime, Protein Boost Immunization Regimen

[0173] This example describes the finding that administration of a c-srRNA vaccine encoding a protein antigen of an original virus is able to prime a humoral immune response to a protein antigen of a variant virus.

Materials and Methods

[0174] BALB/c female mice.

[0175] An srRNA1ts2-G5003 (mRNA), which was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 [WO 2021/138447 A1]) encoding the G5003 antigen (**FIG. 15**).

[0176] Recombinant SARS-CoV-2 B.1.617.2 Spike GCN4-IZ Protein (R&D Systems, Cat. #10878-CV)

[0177] AddaVax™ squalene-based oil-in-water adjuvant was obtained from InvivoGen.

[0178] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through RBD of SARS-CoV-2 (an original Wuhan strain) [JPT Peptides: PepMix SARS-CoV-2 (S-RBD) PM-WCPV-S-RBD-2]

[0179] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0180] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

[0181] Vero76 cells for a plaque reduction neutralization assay (PRNT).

[0182] SARS-CoV-2 Delta Variant live virus for the PRNT assay

[0183] For the PRNT assay, Vero76 cells were first treated with serially diluted mouse serum, followed by the infection with a live virus of SARS-CoV-2 (Delta variant strain). In this assay, the infected cells die and form a plaque after fixation and staining with crystal violet. If the serum contains the neutralizing antibodies, the viral infection is inhibited, resulting in the reduction of the number of plaques. The results are shown as the dilution titer of serum that show 50% reduction of number of plaques ($PRNT_{50}$).

Results

[0184] A composition comprising the c-srRNA encoding G5003 antigen (RBD of SARS-CoV-2 original Wuhan strain) was administered intradermally into skin of BALB/c mice as naked mRNA **(FIG. 17A)**. That is, the srRNA1ts2-G5003 composition did not contain any nanoparticles or transfection reagents. Subsequently, a composition comprising the Spike protein of SARS-CoV-2 (Delta variant B.1.617.2) mixed with adjuvant was administered intradermally (**FIG. 17A**).

[0185] Cellular immunity against the SARS-CoV-2 RBD protein was detected in mouse splenocytes 14 days after a single intradermal injection of the c-srRNA-G5003 composition (**FIG. 17B**). Subsequent exposure of immunized mice to a spike protein of a different SARS-CoV-2 strain (Delta variant B.1.617.2) induced neutralization antibodies (detected by the PRNT

assay) against the Delta variant of SARS-CoV-2 (**FIG. 17C**) as early as day 7 post-protein antigen exposure. In contrast, mice that did not receive c-srRNA-G5003 encoding the RBD of the SARS-CoV-2 (an original Wuhan strain) did not mount a neutralizing antibody response to the Delta variant of SARS-CoV-2. The early induction of neutralizing antibodies is characteristic of a secondary immune response, indicating that the c-srRNA primed the humoral immune response prior to exposure to the adjuvanted RBD protein.

Conclusion

[0186] The results indicate that the c-srRNA immunogen can induce a potent immune response that is broadly reactive against both the antigen encoded by the c-srRNA and a distinct variant antigen. Thus, the c-srRNA SARS-CoV-2 RBD immunogen is suitable for use in immunization regimens directed against a broad spectrum of SARS-CoV-2 strains.

Example 10. Efficacy of Protein Prime, c-srRNA Boost Immunization Regimen

[0187] This example describes the finding that a c-srRNA vaccine can enhance the antibody titer, when used as a booster vaccine for other vaccines.

Materials and Methods

[0188] C57BL/6 female mice.

[0189] RBD protein (Sino Biological SARS-CoV-2 [2019-nCoV] Spike RBD-His Recombinant Protein, Cat. #40592-V08B)

[0190] AddaVax™ squalene-based oil-in-water adjuvant was obtained from InvivoGen.

[0191] An srRNA1ts2-G5003 (mRNA), which was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 [WO 2021/138447 A1]) encoding the G5003 antigen (**FIG. 15**).

[0192] An srRNA1ts2-G5003o (mRNA), which was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 [WO 2021/138447 A1]) encoding the G5003o antigen (**FIG. 15**).

[0193] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through RBD of SARS-CoV-2 (an original Wuhan strain) [JPT Peptides: PepMix SARS-CoV-2 (S-RBD) PM-WCPV-S-RBD-2]

[0194] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0195] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

[0196] ELISA assay plates (ENZO SARS-CoV-2 IgG ELISA Kit [Cat. # ENZ-KIT170-0001, the plate was coated with SARS-CoV-2 (Wuhan strain) S1 antigen RBD protein]).

Results

[0197] To test a possibility whether c-srRNA vaccine can be used as a booster vaccine, mice were first vaccinated with adjuvanted protein (in this case, RBD of SARS-CoV-2 [an original Wuhan strain]). Fourteen days later (Day 14), the mice were further treated with intradermal injection of a placebo (PBO: buffer only), c-srRNA encoding G5003 antigen, c-srRNA encoding G5003o antigen, or the adjuvanted RBD protein (**FIG. 18A**).

[0198] On Day 28, cellular immunity was assessed by the ELISpot assay. As expected, the RBD (1^{st}) + PBO (2^{nd}) group could not induce the cellular immunity, whereas the RBD (1^{st}) + RBD (2^{nd}) group induced the cellular immunity (**FIG. 18B, C**). Interestingly, the RBD (1^{st}) + c-srRNA-G5003 and c-srRNA-G5003o groups also induced the cellular immunity (**FIG. 18B, C**). This was expected, as c-srRNA vaccine alone can induce the cellular immunity.

[0199] On Day 28, the levels of serum antibodies against the RBD of the SARS-CoV-2 virus (an original Wuhan strain) was assessed by an ELISA assay (**FIG. 19**). The first vaccination with the adjuvanted RBD protein alone could induce the antibody weakly. On the other hand, c-srRNA vaccines was able to induce the antibodies at the level, similar to that by the second vaccination with the adjuvanted protein.

Conclusion

[0200] The results indicate that the c-srRNA vaccine can work as a booster vaccine for both cellular immunity and humoral immunity.

Example 11. Potent cellular immune response induced by srRNA1ts2-G5006

[0201] This example describes the finding that a fusion protein comprising the SARS-CoV-2 nucleoprotein and the MERS-CoV nucleoprotein can induce strong cellular immunity against SARS-CoV-2 and MERS-CoV when the protein is expressed from intradermally-

injected, temperature-controllable, self-replicating RNA. The vaccinated mice can eliminate the implanted tumor cells expressing a fusion protein comprising the SARS-CoV-2 nucleoprotein and the MERS-CoV nucleoprotein.

Materials and Methods

[0202] BALB/c female mice.

[0203] srRNA1ts2-G5006 mRNA was produced by in vitro transcription of a temperature-controllable, self-replicating, RNA vector (srRNA1ts2 as described in WO 2021/138447 A1) encoding the G5006 antigen (**FIG. 2**).

[0204] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein (UniProt: P0DTC9) of SARS-CoV-2 [JPT peptide Product Code: PM-WCPV-NCAP].

[0205] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein of MERS-CoV (YP_009047211.1). The peptides were custom-made by JPT Peptides.

[0206] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0207] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

[0208] 4T1 breast cancer cell line, derived from BALB/c mouse and known as a model for a triple-negative stage IV human breast cancer, was purchased from ATCC (catalog # CRL-2539).

[0209] A plasmid DNA, encoding a fusion protein of nucleoproteins of SARS-CoV-2 and MERS-CoV (non-secreted form of G5006, i.e., without CD5 signal peptides) under the CMV promoter, and hygromycin-resistant gene under the promoter of SV40 early promoter, was transfected to 4T1 cells. Cells, expressing the fusion protein of nucleoproteins of SARS-CoV-2 and MERS-CoV (called 4T1-SMN), were isolated by culturing the cells in the presence of 200 μg/mL of hygromycin B.

Results

[0210] To model cells infected with a virus, we used a 4T1 breast cancer cell line, derived from BALB/c mouse and known as a model for a triple-negative stage IV human breast cancer. When injected into BALB/c mouse, the 4T1 cells grow rapidly and form tumors. This syngenic mouse model was used to mimic the rapid increase of infected cells. To this end, we first made a plasmid vector encoding a fusion protein of nucleoproteins of SARS-CoV-2 and MERS-CoV (named SMN protein), under the CMV promoter, so that the protein is constitutively expressed. This fusion protein is the same as G5006, but the CD5 signal peptides were removed from the N-terminus of the protein. Naturally, nucleoprotein does not have the signal peptides and stays within the cytoplasm of the cells. The 4T1 cells expressing the SMN protein (named 4T1-SMN) was established after the hygromycin selection, as the plasmid vector also carried the hygromycin-resistant gene.

[0211] BALB/c mice were vaccinated with c-srRNA-G5006, and the induction of cellular immunity was demonstrated by the presence of T-cells that responded to both SARS-CoV-2 nucleoprotein (**FIG. 20A**) and MERS-CoV nucleoprotein (**FIG. 20B**).

[0212] 4T1-SMN cells were injected into the BALB/c mice vaccinated with c-srRNA-G5006 on day 24 (24 days post-vaccination) (**FIG. 21**). As expected, 4T1-SMN cells grew rapidly mice that received a placebo (no vaccination group) 4T1-SMN cells. On the other hand, the growth of 4T1-SMN tumors were suppressed in the c-srRNA-G5006 vaccinated mice. Two mice received 25 μg of the c-srRNA-G5006 vaccine, though the tumor grew initially, became tumor-free and survived. Furthermore, even after the second round of injection of 4T1-SMN tumors on day 143 after the vaccination, no tumors grew, and the mice were tumor-free and continued to live (**FIG. 21**).

Conclusion

[0213] c-srRNA vaccine can induce strong cellular immunity, which can kill and eliminating cells that express the antigen. This result indicates that c-srRNA functions as a vaccine by eliminating the infected cells.

Example 12. Cellular immunity induced by srRNA1ts2-PanCoronavirus Vaccine

[0214] This example describes the finding that a fusion protein comprising the CD5 signal peptides, Spike-RBD of SARS-CoV-2, nucleoprotein of SARS-CoV-2, nucleoprotein of MERS-CoV, and Spike-RBD of MERS-CoV can induce strong cellular immunity against all of these antigens, when the protein is expressed from intradermally-injected, temperature-controllable, self-replicating RNA.

Materials and Methods

[0215] C57BL/6 female mice.

[0216] srRNA1ts2-G5006 mRNA was produced by in vitro transcription of a temperature-controllable, self-replicating, RNA vector (srRNA1ts2 as described in WO 2021/138447 A1) encoding the G5006 antigen (**FIG. 2**).

[0217] srRNA1ts2-G5006d mRNA was produced by in vitro transcription of a temperature-controllable, self-replicating, RNA vector (srRNA1ts2 as described in WO 2021/138447 A1) encoding the G5006d antigen (**FIG. 22**).

[0218] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through RBD of SARS-CoV-2 (an original Wuhan strain) [JPT Peptides: PepMix SARS-CoV-2 (S-RBD) PM-WCPV-S-RBD-2]

[0219] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein (UniProt: P0DTC9) of SARS-CoV-2 [JPT peptide Product Code: PM-WCPV-NCAP].

[0220] A pool of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Nucleoprotein of MERS-CoV (YP_009047211.1). The peptides were custom-made by JPT Peptides.

[0221] A pool of 336 (168+168) peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through Spike glycoprotein (Swiss-Prot ID: K9N5Q8) of MERS-CoV (Middle East respiratory syndrome-related coronavirus) [JPT peptides Product Code: PM-MERS-CoV-S-1].

[0222] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0223] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

Results

[0224] We here designed a new booster vaccine, which is a c-srRNA vaccine (called c-srRNA-G5006d) encoding a fusion protein comprising the CD5 signal peptides, Spike-RBD of SARS-CoV-2, nucleoprotein of SARS-CoV-2, nucleoprotein of MERS-CoV, and Spike-RBD of MERS-CoV (**FIG. 22**).

[0225] Mice were vaccinated with the intradermal injection of a placebo (PBO: buffer only), c-srRNA encoding G5006 antigen, and c-srRNA encoding G5006d antigen. On day 14 post-vaccination, cellular immunity was assessed by ELISpot assays.

[0226] As shown in **FIG. 23A**, c-srRNA-G5006d can stimulate cellular immunity against all the proteins encoded on this vaccine: Spike-RBD of SARS-CoV-2, Nucleoprotein of SARS-CoV-2, Nucleoprotein of MERS-CoV, and Spike-RBD of MERS-CoV.

Conclusion

[0227] The results indicate that the c-srRNA vaccine can work as a booster vaccine for both cellular immunity and humoral immunity.

Example 13. srRNA1ts2-PanInfluenza Virus Vaccine

[0228] This example describes the design of pan-influenza booster vaccine based on the unique feature of c-srRNA vaccine platform. An antigen (G5012) encoded on c-srRNA is a fusion protein of CD5 signal peptide (residues 1-24), a part of the hemagglutinin (HA) of the Influenza A, nucleoprotein of Influenza A, nucleoprotein of Influenza B, and a part of the hemagglutinin (HA) of the Influenza B.

Materials and Methods

[0229] C57BL/6 female mice.

[0230] c-srRNA-G5012 mRNA was produced by in vitro transcription of a temperature-controllable, self-replicating, RNA vector (srRNA1ts2 as described in WO 2021/138447 A1) encoding the G5012 antigen (**FIG. 24**).

[0231] Pools of peptides derived from a peptide scan (15mers with 11 amino acid overlaps) through a part of the hemagglutinin (HA) of the Influenza A, nucleoprotein of Influenza A, nucleoprotein of Influenza B, and a part of the hemagglutinin (HA) of the Influenza B.

[0232] ELISpot assay plates and reagents for interferon gamma (INF-γ) and interleukin-4 (IL-4) (Cellular Technology Limited, Ohio, USA).

[0233] Immunospot S6 Entry Analyzer (Cellular Technology Limited, Ohio, USA).

Results

[0234] Mice were vaccinated with the intradermal injection of a placebo (PBO: buffer only), and c-srRNA encoding G5012 antigen. On day 14 post-vaccination, cellular immunity was assessed by ELISpot assays.

[0235] c-srRNA-G5012 stimulated cellular immunity against all the antigen encoded on this vaccine: the hemagglutinin (HA) of the Influenza A, the nucleoprotein of Influenza A, the nucleoprotein of Influenza B, and the hemagglutinin (HA) of the Influenza B.

Conclusion

[0236] The results indicate that the c-srRNA vaccine can work as a booster vaccine for both cellular immunity and humoral immunity.

Example 14. Chitosan-enhanced luciferase expression from srRNA1ts2-LUC2

[0237] This example describes the finding that chitosan oligomers are able to enhance in vivo expression of a gene of interest (GOI) encoded by a c-srRNA construct.

Materials and Methods

[0238] C57BL/6 female mice.

[0239] An srRNA1ts2-LUC2 (mRNA), which was produced by in vitro transcription of a temperature-controllable self-replicating RNA vector (srRNA1ts2 as described in WO 2021/138447 A1) encoding the luciferase gene.

[0240] Chitosan Oligomer (molecular weight ≤ 5 kDa, ≥75.0% deacetylated: Heppe Medical Chitosan GmbH: Product No. 44009)

[0241] Chitosan oligosaccharide lactate (molecular weight ~ 5 kDa, > 90% deacetylated: Sigma-Aldrich: Product No. 523682)

[0242] Bioluminescent Imaging system, AMI HTX (Spectral Instruments Imaging, Tucson, AZ)

Results

[0243] To test whether chitosan oligomers can enhance the expression of GOI encoded on c-srRNA in vivo, 5 μg of c-srRNA (also known as srRNA1ts2) encoding a luciferase gene as GOI was mixed with chitosans and administered intradermally to each C57BL/6 mouse (**FIG. 25**). c-srRNAs were formulated as naked RNAs, without lipid nanoparticles or any other transfection reagents, in lactated Ringer's solution. Luciferase activity was visualized and quantitated by using a bioluminescent Imaging system, AMI HTX (Spectral Instruments Imaging, Tucson, AZ).

[0244] Five mice each were tested in the following groups: 1, a control – c-srRNA only; 2, c-srRNA mixed with chitosan oligosaccharide (0.001 μg/mL); 3, c-srRNA mixed with chitosan oligosaccharide (0.01 μg/mL); 4, c-srRNA mixed with chitosan oligosaccharide (0.5 μg/mL); 5, c-srRNA mixed with chitosan oligosaccharide lactate (0.1 μg/mL).

[0245] As shown in FIG. 27, compared to the control condition (i.e., c-srRNA only: no chitosan), all the conditions with chitosan oligomers at the concentration of 0.001 μg/mL, 0.01 μg/mL, and 0.5 μg/mL as well as the condition with chitosan oligosaccharide lactate at the concentration of 0.1 μg/mL showed ~10-fold higher levels of luciferase activity.

Conclusion

[0246] Low-molecular-weight chitosans such as chitosan oligomers and chitosan oligosaccharide lactate can enhance the expression of GOI encoded on c-srRNA, when mixed

with c-srRNA before injecting c-srRNA into mouse skin intradermally. Chitosan oligomers provide about a 10-fold enhancement of gene expression even at a very low concentration (0.001 µg/mL or about 0.2 nM). This surprising discovery provides an effective means to enhance the in vivo therapeutic expression of GOI encoded on c-srRNA.

SEQUENCES

SEQ ID NO:1
>G5004 (SARS-COV-2_N, mRNA)
ATGTCTGATAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACCCTCAG
ATTCAACTGGCAGTAACCAGAATGGAGAACGCAGTGGGGCGCGATCAAAACAACGTCGGCCCCA
AGGTTTACCCAATAATACTGCGTCTTGGTTCACCGCTCTCACTCAACATGGCAAGGAAGACCTT
AAATTCCCTCGAGGACAAGGCGTTCCAATTAACACCAATAGCAGTCCAGATGACCAAATTGGCT
ACTACCGAAGAGCTACCAGACGAATTCGTGGTGGTGACGGTAAAATGAAAGATCTCAGTCCAAG
ATGGTATTTCTACTACCTAGGAACTGGGCCAGAAGCTGGACTTCCCTATGGTGCTAACAAAGAC
GGCATCATATGGGTTGCAACTGAGGGAGCCTTGAATACACCAAAAGATCACATTGGCACCCGCA
ATCCTGCTAACAATGCTGCAATCGTGCTACAACTTCCTCAAGGAACAACATTGCCAAAAGGCTT
CTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTCCTCATCACGTAGTCGCAAC
AGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACTTCTCCTGCTAGAATGGCTGGCAATG
GCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGCTTGAGAGCAAAATGTC
TGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAAGAAATCTGCTGCTGAGGCTTCTAAG
AAGCCTCGGCAAAAACGTACTGCCACTAAAGCATACAATGTAACACAAGCTTTCGGCAGACGTG
GTCCAGAACAAACCCAAGGAAATTTTGGGGACCAGGAACTAATCAGACAAGGAACTGATTACAA
ACATTGGCCGCAAATTGCACAATTTGCCCCCAGCGCTTCAGCGTTCTTCGGAATGTCGCGCATT
GGCATGGAAGTCACACCTTCGGGAACGTGGTTGACCTACACAGGTGCCATCAAATTGGATGACA
AAGATCCAAATTTCAAAGATCAAGTCATTTTGCTGAATAAGCATATTGACGCATACAAAACATT
CCCACCAACAGAGCCTAAAAAGGACAAAAAGAAGAAGGCTGATGAAACTCAAGCCTTACCGCAG
AGACAGAAGAAACAGCAAACTGTGACTCTTCTTCCTGCTGCAGATTTGGATGATTTCTCCAAAC
AATTGCAACAATCCATGAGCAGTGCTGACTCAACTCAGGCCTAA

SEQ ID NO:2
>G5005 (CD5-SP_SARS-COV-2_N, synthetic mRNA)
ATGCCCATGGGGTCTCTGCAACCGCTGGCCACCTTGTACCTGCTGGGGATGCTGGTCGCTTCCT
GCCTCGGATCTGATAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACC
CTCAGATTCAACTGGCAGTAACCAGAATGGAGAACGCAGTGGGGCGCGATCAAAACAACGTCGG
CCCCAAGGTTTACCCAATAATACTGCGTCTTGGTTCACCGCTCTCACTCAACATGGCAAGGAAG
ACCTTAAATTCCCTCGAGGACAAGGCGTTCCAATTAACACCAATAGCAGTCCAGATGACCAAAT
TGGCTACTACCGAAGAGCTACCAGACGAATTCGTGGTGGTGACGGTAAAATGAAAGATCTCAGT
CCAAGATGGTATTTCTACTACCTAGGAACTGGGCCAGAAGCTGGACTTCCCTATGGTGCTAACA
AAGACGGCATCATATGGGTTGCAACTGAGGGAGCCTTGAATACACCAAAAGATCACATTGGCAC
CCGCAATCCTGCTAACAATGCTGCAATCGTGCTACAACTTCCTCAAGGAACAACATTGCCAAAA
GGCTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTCCTCATCACGTAGTC
GCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACTTCTCCTGCTAGAATGGCTGG
CAATGGCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGCTTGAGAGCAAA

ATGTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAAGAAATCTGCTGCTGAGGCTT
CTAAGAAGCCTCGGCAAAAACGTACTGCCACTAAAGCATACAATGTAACACAAGCTTTCGGCAG
ACGTGGTCCAGAACAAACCCAAGGAAATTTTGGGGACCAGGAACTAATCAGACAAGGAACTGAT
TACAAACATTGGCCGCAAATTGCACAATTTGCCCCCAGCGCTTCAGCGTTCTTCGGAATGTCGC
GCATTGGCATGGAAGTCACACCTTCGGGAACGTGGTTGACCTACACAGGTGCCATCAAATTGGA
TGACAAAGATCCAAATTTCAAAGATCAAGTCATTTTGCTGAATAAGCATATTGACGCATACAAA
ACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAGAAGGCTGATGAAACTCAAGCCTTAC
CGCAGAGACAGAAGAAACAGCAAACTGTGACTCTTCTTCCTGCTGCAGATTTGGATGATTTCTC
CAAACAATTGCAACAATCCATGAGCAGTGCTGACTCAACTCAGGCCTAA

SEQ ID NO:3
>G5006A (CD5-SP_SARS2-N_MERS-N, synthetic mRNA)
ATGCCCATGGGGTCTCTGCAACCGCTGGCCACCTTGTACCTGCTGGGGATGCTGGTCGCTTCCT
GCCTCGGATCTGATAATGGACCCCAAAATCAGCGAAATGCACCCCGCATTACGTTTGGTGGACC
CTCAGATTCAACTGGCAGTAACCAGAATGGAGAACGCAGTGGGGCGCGATCAAAACAACGTCGG
CCCCAAGGTTTACCCAATAATACTGCGTCTTGGTTCACCGCTCTCACTCAACATGGCAAGGAAG
ACCTTAAATTCCCTCGAGGACAAGGCGTTCCAATTAACACCAATAGCAGTCCAGATGACCAAAT
TGGCTACTACCGAAGAGCTACCAGACGAATTCGTGGTGGTGACGGTAAAATGAAAGATCTCAGT
CCAAGATGGTATTTCTACTACCTAGGAACTGGGCCAGAAGCTGGACTTCCCTATGGTGCTAACA
AAGACGGCATCATATGGGTTGCAACTGAGGGAGCCTTGAATACACCAAAAGATCACATTGGCAC
CCGCAATCCTGCTAACAATGCTGCAATCGTGCTACAACTTCCTCAAGGAACAACATTGCCAAAA
GGCTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTCCTCATCACGTAGTC
GCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACTTCTCCTGCTAGAATGGCTGG
CAATGGCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGCTTGAGAGCAAA
ATGTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAAGAAATCTGCTGCTGAGGCTT
CTAAGAAGCCTCGGCAAAAACGTACTGCCACTAAAGCATACAATGTAACACAAGCTTTCGGCAG
ACGTGGTCCAGAACAAACCCAAGGAAATTTTGGGGACCAGGAACTAATCAGACAAGGAACTGAT
TACAAACATTGGCCGCAAATTGCACAATTTGCCCCCAGCGCTTCAGCGTTCTTCGGAATGTCGC
GCATTGGCATGGAAGTCACACCTTCGGGAACGTGGTTGACCTACACAGGTGCCATCAAATTGGA
TGACAAAGATCCAAATTTCAAAGATCAAGTCATTTTGCTGAATAAGCATATTGACGCATACAAA
ACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAGAAGGCTGATGAAACTCAAGCCTTAC
CGCAGAGACAGAAGAAACAGCAAACTGTGACTCTTCTTCCTGCTGCAGATTTGGATGATTTCTC
CAAACAATTGCAACAATCCATGAGCAGTGCTGACTCAACTCAGGCCATGGCATCCCCTGCTGCA
CCTCGTGCTGTTTCCTTTGCCGATAACAATGATATAACAAATACAAACCTATCTCGAGGTAGAG
GACGTAATCCAAAACCACGAGCTGCACCAAATAACACTGTCTCTTGGTACACTGGGCTTACCCA
ACACGGGAAAGTCCCTCTTACCTTTCCACCTGGGCAGGGTGTACCTCTTAATGCCAATTCTACC
CCTGCGCAAAATGCTGGGTATTGGCGGAGACAGGACAGAAAAATTAATACCGGGAATGGAATTA
AGCAACTGGCTCCCAGGTGGTACTTCTACTACACTGGAACTGGACCCGAAGCAGCACTCCCATT
CCGGGCTGTTAAGGATGGCATCGTTTGGGTCCATGAAGATGGCGCCACTGATGCTCCTTCAACT
TTTGGGACGCGGAACCCTAACAATGATTCAGCTATTGTTACACAATTCGCGCCCGGTACTAAGC
TTCCTAAAAACTTCCACATTGAGGGGACTGGAGGCAATAGTCAATCATCTTCAAGAGCCTCTAG
CTTAAGCAGAAACTCTTCCAGATCTAGTTCACAAGGTTCAAGATCAGGAAACTCTACCCGCGGC
ACTTCTCCAGGTCCATCTGGAATCGGAGCAGTAGGAGGTGATCTACTTTACCTTGATCTTCTGA
ACAGACTACAAGCCCTTGAGTCTGGCAAAGTAAAGCAATCGCAGCCAAAAGTAATCACTAAGAA
AGATGCTGCTGCTGCTAAAAATAAGATGCGCCACAAGCGCACTTCCACCAAAAGTTTCAACATG
GTGCAAGCTTTTGGTCTTCGCGGACCAGGAGACCTCCAGGGAAACTTTGGTGATCTTCAATTGA
ATAAACTCGGCACTGAGGACCCACGTTGGCCCCAAATTGCTGAGCTTGCTCCTACAGCCAGTGC

TTTTATGGGTATGTCGCAATTTAAACTTACCCATCAGAACAATGATGATCATGGCAACCCTGTG
TACTTCCTTCGGTACAGTGGAGCCATTAAACTTGACCCAAAGAATCCCAACTACAATAAGTGGT
TGGAGCTTCTTGAGCAAAATATTGATGCCTACAAAACCTTCCCTAAGAAGGAAAAGAAACAAAA
GGCACCAAAAGAAGAATCAACAGACCAAATGTCTGAACCTCCAAAGGAGCAGCGTGTGCAAGGT
AGCATCACTCAGCGCACTCGCACCCGTCCAAGTGTTCAGCCTGGTCCAATGATTGATGTTAACA
CTGATTAG

SEQ ID NO:4
>G5006B (CD5-SP_SARS2-N_MERS-N: CODON-OPTIMIZED, synthetic mRNA)
ATGCCTATGGGCTCTCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCAGCT
GCCTGGGAAGCGACAACGGCCCCCAGAACCAGAGAAACGCCCCTAGAATCACATTTGGCGGCCC
TAGTGATAGCACCGGATCTAATCAAAACGGCGAGAGAAGCGGCGCTCGGTCTAAACAGAGACGG
CCACAGGGACTGCCTAACAACACCGCCAGCTGGTTCACCGCCCTGACCCAGCACGGCAAGGAGG
ACCTTAAGTTCCCTCGGGGACAGGGCGTGCCAATCAACACCAACTCTAGTCCCGACGACCAGAT
CGGCTATTATAGAAGAGCCACAAGACGCATCAGAGGTGGCGACGGCAAGATGAAGGACCTGAGC
CCTCGCTGGTACTTTTACTACCTGGGGACCGGCCCTGAAGCCGGCCTGCCTTACGGCGCCAACA
AGGACGGAATCATCTGGGTCGCCACCGAGGGCGCCCTGAATACCCCTAAGGACCACATCGGCAC
CAGAAACCCTGCTAATAATGCCGCTATCGTGCTGCAGCTGCCTCAGGGCACCACCCTGCCTAAG
GGCTTCTACGCCGAGGGCTCCCGGGGAGGTTCCCAGGCTAGCAGCAGATCTTCCAGCCGGAGCA
GAAACAGCTCCAGGAACAGCACACCTGGCAGCAGCAGAGGTACGAGCCCTGCCCGGATGGCCGG
AAACGGCGGCGATGCCGCCCTGGCCCTGCTGCTGCTGGACAGACTGAACCAGCTCGAGAGCAAG
ATGTCTGGCAAGGGCCAGCAGCAGCAGGGCCAGACAGTGACCAAGAAATCCGCCGCTGAGGCCA
GCAAAAAACCCAGACAGAAAAGAACCGCTACAAAGGCCTACAACGTTACCCAGGCCTTTGGCAG
ACGGGGCCCAGAGCAGACCCAGGGAAACTTCGGCGACCAGGAGCTGATCCGGCAGGGCACGGAC
TACAAGCACTGGCCTCAAATCGCCCAGTTTGCCCCTTCCGCCAGCGCTTTCTTCGGAATGAGCA
GAATCGGCATGGAAGTGACACCTAGTGGCACCTGGCTGACCTACACCGGTGCCATTAAGCTGGA
TGACAAGGACCCCAACTTCAAGGATCAGGTGATCCTGCTGAACAAGCACATTGATGCTTACAAG
ACCTTCCCACCTACCGAGCCAAAAAAAGATAAGAAGAAAAAAGCCGATGAGACACAAGCCCTGC
CCCAGAGGCAGAAGAAGCAACAAACCGTCACCCTGCTGCCTGCTGCCGACCTGGACGACTTCAG
CAAACAGCTGCAGCAGAGCATGAGCTCTGCTGATAGCACCCAGGCCATGGCCTCTCCAGCCGCT
CCCAGAGCTGTGTCCTTCGCCGATAATAACGACATCACAAACACCAACCTGAGCCGGGGCAGAG
GCAGAAACCCTAAACCTAGAGCCGCCCCCAACAACACCGTGAGCTGGTATACAGGCCTCACCCA
GCATGGCAAGGTGCCTCTGACATTCCCCCCTGGGCAGGGCGTGCCCCTGAACGCCAACAGCACC
CCTGCCCAGAATGCCGGCTACTGGCGGAGGCAGGACAGAAAGATCAACACTGGTAACGGCATCA
AGCAGCTGGCCCCACGGTGGTATTTCTACTACACCGGCACCGGCCCGGAAGCCGCCCTCCCCTT
CAGAGCCGTGAAGGACGGCATCGTGTGGGTGCACGAGGACGGCGCCACAGATGCCCCGTCTACA
TTTGGCACTCGGAATCCCAATAACGACAGCGCCATCGTGACCCAGTTCGCCCCTGGCACCAAGC
TGCCTAAGAACTTTCACATCGAGGGCACAGGAGGCAACAGCCAGAGCAGCAGCCGGGCTTCGAG
CCTGTCTCGGAATAGCTCCCGGTCCAGCTCTCAGGGCAGCCGCAGTGGAAATTCCACCCGGGGC
ACATCTCCTGGCCCCAGCGGCATCGGCGCTGTGGGCGGAGACCTGCTCTACCTGGACCTGCTGA
ACAGACTGCAGGCACTTGAAAGCGGCAAAGTTAAGCAATCTCAACCTAAGGTGATCACCAAAAA
GGACGCCGCCGCCGCTAAGAACAAGATGAGACACAAGAGAACAAGCACAAAGAGCTTCAACATG
GTGCAAGCCTTCGGCCTGCGGGGACCTGGCGACCTGCAGGGCAACTTCGGCGACCTGCAGCTGA
ACAAGCTGGGCACAGAGGATCCTCGATGGCCCCAGATCGCCGAACTAGCTCCAACCGCCAGCGC
CTTCATGGGCATGAGCCAGTTCAAGCTGACACACCAGAACAATGACGATCACGGAAATCCTGTG
TACTTCCTGAGATACAGCGGCGCCATCAAGCTGGATCCTAAGAACCCCAACTACAACAAGTGGC
TGGAACTGCTGGAACAGAACATCGACGCCTACAAGACCTTCCCCAAGAAGGAAAAGAAGCAGAA

GGCCCCTAAAGAGGAAAGCACAGATCAGATGAGCGAGCCTCCCAAGGAACAGAGAGTGCAGGGA
TCTATCACCCAGAGAACAAGAACAAGACCCAGCGTGCAGCCTGGCCCTATGATTGACGTGAACA
CCGACTAG

SEQ ID NO:5
>G5004 (SARS-COV-2_N, protein)
MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQGLPNNTASWFTALTQHGKEDL
KFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPYGANKD
GIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAEGSRGGSQASSRSSSRSRN
SSRNSTPGSSRGTSPARMAGNGGDAALALLLLDRLNQLESKMSGKGQQQQGQTVTKKSAAEASK
KPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRI
GMEVTPSGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQ
RQKKQQTVTLLPAADLDDFSKQLQQSMSSADSTQA

SEQ ID NO:6
>G5005 (CD5-SP_SARS-CoV-2_N, synthetic protein)
MPMGSLQPLATLYLLGMLVASCLGSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRR
PQGLPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLS
PRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPK
GFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGDAALALLLLDRLNQLESK
MSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTD
YKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYK
TFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAADLDDFSKQLQQSMSSADSTQA

SEQ ID NO:7
>G5006a (CD5-SP_SARS2-N_MERS-N, synthetic protein)
MPMGSLQPLATLYLLGMLVASCLGSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRR
PQGLPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLS
PRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPK
GFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGDAALALLLLDRLNQLESK
MSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTD
YKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYK
TFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAADLDDFSKQLQQSMSSADSTQA
MASPAAPRAVSFADNNDITNTNLSRGRGRNPKPRAAPNNTVSWYTGLTQHGKVPLTFPPGQGVP
LNANSTPAQNAGYWRRQDRKINTGNGIKQLAPRWYFYYTGTGPEAALPFRAVKDGIVWVHEDGA
TDAPSTFGTRNPNNDSAIVTQFAPGTKLPKNFHIEGTGGNSQSSSRASSLSRNSSRSSSQGSRS
GNSTRGTSPGPSGIGAVGGDLLYLDLLNRLQALESGKVKQSQPKVITKKDAAAAKNKMRHKRTS
TKSFNMVQAFGLRGPGDLQGNFGDLQLNKLGTEDPRWPQIAELAPTASAFMGMSQFKLTHQNND
DHGNPVYFLRYSGAIKLDPKNPNYNKWLELLEQNIDAYKTFPKKEKKQKAPKEESTDQMSEPPK
EQRVQGSITQRTRTRPSVQPGPMIDVNTD

SEQ ID NO:8
>human CD5 Signal Peptide (Homo sapiens, protein)
MPMGSLQPLATLYLLGMLVASCLG

SEQ ID NO:9
>MERS-N protein

MASPAAPRAVSFADNNDITNTNLSRGRGRNPKPRAAPNNTVSWYTGLTQHGKVPLTFPPGQGVP
LNANSTPAQNAGYWRRQDRKINTGNGIKQLAPRWYFYYTGTGPEAALPFRAVKDGIVWVHEDGA
TDAPSTFGTRNPNNDSAIVTQFAPGTKLPKNFHIEGTGGNSQSSSRASSLSRNSSRSSSQGSRS
GNSTRGTSPGPSGIGAVGGDLLYLDLLNRLQALESGKVKQSQPKVITKKDAAAAKNKMRHKRTS
TKSFNMVQAFGLRGPGDLQGNFGDLQLNKLGTEDPRWPQIAELAPTASAFMGMSQFKLTHQNND
DHGNPVYFLRYSGAIKLDPKNPNYNKWLELLEQNIDAYKTFPKKEKKQKAPKEESTDQMSEPPK
EQRVQGSITQRTRTRPSVQPGPMIDVNTD

SEQ ID NO:10
>SARS-COV-1-N protein
MSDNGPQSNQRSAPRITFGGPTDSTDNNQNGGRNGARPKQRRPQGLPNNTASWFTALTQHGKEE
LRFPRGQGVPINTNSGPDDQIGYYRRATRRVRGGDGKMKELSPRWYFYYLGTGPEASLPYGANK
EGIVWVATEGALNTPKDHIGTRNPNNNAATVLQLPQGTTLPKGFYAEGSRGGSQASSRSSSRSR
GNSRNSTPGSSRGNSPARMASGGGETALALLLLDRLNQLESKVSGKGQQQQGQTVTKKSAAEAS
KKPRQKRTATKQYNVTQAFGRRGPEQTQGNFGDQDLIRQGTDYKHWPQIAQFAPSASAFFGMSR
IGMEVTPSGTWLTYHGAIKLDDKDPQFKDNVILLNKHIDAYKTFPPTEPKKDKKKKTDEAQPLP
QRQKKQPTVTLLPAADMDDFSRQLQNSMSGASADSTQA

SEQ ID NO:11
>cloning site (synthetic DNA)
GGCGCGCCattggccaccGCGGCCGC
The gene of interest (e.g., betacoronavirus nucleoprotein open reading frame) is inserted between nucleotides 18 and 19 of SEQ ID NO:11. GGCGCGCC is an AscI restriction site and GCGGCCGC is a NotI restriction site.

SEQ ID NO:12
>artificial
Attggccacc (synthetic DNA)
This nucleotide sequence is added to the 5′ end of the betacoronavirus nucleoprotein coding region to provide a Kozak consensus sequence for initiation of translation of the mRNA.

SEQ ID NO:13
>Influenza A virus (H5N8) nucleocapsid protein
ASQGTKRSYEQMETGGERQNATEIRASVGRMVGGIGRFYIQMCTELKLSDYEGRLIQNSITIER
MVLSAFDERRNKYLEEHPSAGKDPKKTGGPIYRRRDGKWVRELILYDKEEIRRIWRQANNGEDA
TAGLTHLMIWHSNLNDATYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGVGTMVMEL
IRMIKRGINDRNFWRGENGRRTRIAYERMCNILKGKFQTAAQRAMMDQVRESRNPGNAEIEDLI
FLARSALILRGSVAHKSCLPACVYGLAVASGYDFEREGYSLVGIDPFRLLQNSQVFSLIRPNEN
PAHKSQLVWMACHSAAFEDLRVSSFIRGTRVVPRGQLSTRGVQIASNENMETMDSSTLELRSRY
WAIRTRSGGTTNQQRASAGQISVQPTFSVQRNLPFERATIMAAFTGNTEGRTSDMRTEIIRMME
SAKPEDVSFQGRGVFELSDEKATNPIVPSFDMSNEGSYFFGDNAEEYDN
>
SEQ ID NO:14
>Influenza B virus nucleocapsid protein

MSNMDIDGINTGTIDKTPEEITPGTSGTTRPIIRPATLAPPSNKRTRNPSPERATTSSEDDVGR
KTQKKQTPTEIKKSVYNMVVKLGEFYNQMMVKAGLNDDMERNLIQNAHAVERILLAATDDKKTE
FQKKKNARDVKEGKEEIDHNKTGGTFYKMVRDDKTIYFSPIRITFLKEEVKTMYKTTMGSDGFS
GLNHIMIGHSQMNDVCFQRSKALKRVGLDPSLISTFAGSTIPRRSGATGVAIKGGGTLVAEAIR
FIGRAMADRGLLRDIKAKTAYEKILLNLKNKCSAPQQKALVDQVIGSRNPGIADIEDLTLLARS
MVVVRPSVASKVVLPISIYAKIPQLGFNVEEYSMVGYEAMALYNMATPVSILRMGDDAKDKSQL
FFMSCFGAAYEDLRVLSALTGTEFKPRSALKCKGFHVPAKEQVEGMGAALMSIKLQFWAPMTRS
GGNEVGGDGGSGQISCSPVFAVERPIALSKQAVRRMLSMNIEGRDADVKGNLLKMMNDSMAKKT
SGNAFIGKKMFQISDKNKTNPVEIPIKQTIPNFFFGRDTAEDYDDLDY

SEQ ID NO:15
>G5010 (synthetic mRNA) [codon-optimized_sequence_for_Pan-Influenza_Antigen]
ATGCCTATGGGCAGCCTGCAGCCACTGGCTACACTGTACCTGCTGGGCATGCTGGTGGCCTCTT
GTCTGGGCGCCAGCCAAGGCACTAAGAGAAGCTACGAGCAGATGGAAACCGGAGGCGAACGGCA
GAACGCCACAGAGATCAGAGCCTCTGTGGGCCGTATGGTCGGCGGCATCGGCAGATTCTACATC
CAGATGTGCACCGAACTGAAGCTGAGCGACTACGAGGGCCGCCTGATCCAGAACAGCATCACAA
TCGAGAGAATGGTGCTGTCCGCCTTTGACGAGCGGAGAAACAAATACCTGGAAGAGCACCCTAG
CGCCGGAAAAGATCCTAAGAAAACCGGCGGACCTATCTACAGAAGAAGAGATGGTAAGTGGGTG
AGAGAGCTGATTCTGTACGATAAGGAAGAGATTCGAAGAATCTGGAGACAGGCCAACAACGGCG
AGGATGCCACCGCAGGCCTGACACACCTGATGATCTGGCACAGCAACCTGAACGATGCGACCTA
CCAGCGCACGCGGGCCCTGGTCAGAACCGGCATGGATCCTCGGATGTGTAGCCTGATGCAGGGC
AGCACACTGCCAAGACGGAGTGGGGCCGCCGGCGCTGCAGTGAAGGGCGTCGGAACCATGGTGA
TGGAGCTGATCCGGATGATAAAGCGGGGCATCAACGACAGAAACTTCTGGCGAGGCGAGAACGG
CCGAAGAACCCGGATCGCCTACGAGAGAATGTGCAACATCCTGAAAGGAAAATTCCAGACCGCC
GCCCAGCGGGCCATGATGGACCAGGTGCGCGAGAGCAGAAACCCCGGCAATGCCGAGATCGAGG
ACCTGATCTTCCTGGCCAGAAGCGCCCTCATTCTTAGAGGCTCTGTGGCCCACAAGAGCTGTCT
GCCTGCCTGTGTGTACGGCCTGGCAGTGGCCTCAGGCTACGACTTCGAGCGGGAAGGATACAGT
CTGGTGGGCATCGACCCTTTCAGACTCCTGCAGAATAGCCAGGTGTTTAGCCTGATCAGACCAA
ACGAAAACCCCGCCCATAAGAGCCAGCTGGTGTGGATGGCCTGCCACAGCGCCGCCTTTGAGGA
TCTGAGAGTGAGCTCTTTTATCAGAGGCACCCGGGTGGTTCCACGAGGTCAACTGTCTACAAGA
GGTGTGCAGATCGCCAGCAACGAGAACATGGAGACCATGGATAGCAGCACCCTGGAACTGAGAT
CCAGATACTGGGCCATCAGGACACGGAGCGGCGGCACCACCAATCAGCAGCGCGCCAGCGCCGG
CCAGATCTCTGTCCAGCCTACGTTTAGCGTGCAGCGGAATTTGCCCTTCGAACGCGCCACAATC
ATGGCTGCTTTCACCGGCAATACAGAGGGCAGAACCAGCGATATGAGAACAGAAATTATCCGTA
TGATGGAGTCCGCAAAACCTGAGGACGTGTCCTTCCAAGGCAGAGGCGTGTTCGAGCTGAGCGA
CGAGAAGGCCACCAACCCTATCGTGCCTAGCTTCGATATGTCTAATGAGGGCAGCTACTTTTTC
GGAGATAACGCCGAAGAGTACGACAACATGTCTAATATGGATATCGACGGCATTAACACCGGCA
CCATCGACAAAACCCCTGAGGAGATCACCCCTGGCACCAGCGGCACAACCCGGCCCATCATCCG
CCCCGCTACACTGGCTCCACCTAGCAACAAGCGGACCAGAAATCCCTCGCCAGAAAGAGCCACA
ACCTCCAGCGAGGACGACGTGGGACGGAAGACACAAAAGAAGCAGACCCCTACAGAGATCAAGA
AGTCTGTTTACAACATGGTGGTGAAACTGGGCGAGTTCTACAACCAGATGATGGTGAAGGCCGG
CCTGAACGACGATATGGAAAGAAATCTGATCCAGAACGCCCACGCCGTGGAGCGGATTCTGCTG
GCCGCCACCGATGATAAGAAGACCGAATTCCAGAAAAAGAAAAACGCCAGAGACGTGAAGGAAG
GCAAGGAAGAGATCGACCACAACAAGACAGGCGGCACATTCTACAAGATGGTCCGGGACGACAA
GACCATCTACTTCAGCCCTATCCGGATAACATTCCTGAAAGAAGAAGTGAAGACCATGTACAAA
ACCACAATGGGCTCTGACGGCTTCAGCGGCCTGAATCACATCATGATCGGCCACTCTCAAATGA

ACGATGTGTGCTTCCAGAGAAGCAAGGCTCTGAAGCGCGTGGGCCTGGATCCTAGCCTGATCTC
TACCTTCGCCGGCAGCACCATCCCCAGAAGATCGGGCGCTACCGGCGTGGCTATCAAGGGAGGA
GGCACACTGGTGGCTGAAGCCATCAGATTCATCGGAAGAGCCATGGCCGACAGAGGACTCCTGA
GAGATATCAAAGCCAAAACCGCCTACGAAAAAATCCTGCTGAACCTGAAGAACAAGTGCAGCGC
GCCTCAACAGAAGGCCCTGGTGGACCAGGTTATCGGCTCTAGAAACCCTGGAATCGCCGATATC
GAGGACCTGACACTGCTGGCCAGATCTATGGTGGTGGTGAGACCCTCCGTGGCCAGCAAGGTGG
TGCTGCCTATCAGCATCTACGCCAAGATCCCTCAGCTGGGATTTAACGTGGAAGAATACAGCAT
GGTTGGTTATGAGGCCATGGCCCTGTACAACATGGCCACACCTGTGTCCATCCTGAGAATGGGC
GACGATGCCAAAGACAAGAGCCAGCTGTTCTTCATGAGCTGCTTCGGCGCTGCCTATGAGGACC
TGAGAGTGCTGTCCGCTCTTACAGGAACAGAGTTCAAGCCTAGGAGCGCACTGAAGTGCAAGGG
CTTCCACGTGCCCGCCAAGGAACAGGTGGAAGGCATGGGAGCTGCTCTGATGTCCATCAAGCTG
CAATTTTGGGCTCCTATGACCCGGAGCGGCGGAAATGAGGTGGGTGGCGACGGAGGCAGCGGAC
AGATTTCTTGCAGCCCCGTATTTGCCGTGGAGAGACCAATCGCCCTGTCCAAGCAGGCCGTGAG
AAGAATGCTGAGCATGAACATCGAGGGCCGGGACGCCGACGTGAAGGGCAACCTGTTGAAGATG
ATGAACGACAGCATGGCCAAGAAGACCAGTGGCAATGCCTTCATCGGCAAGAAGATGTTCCAGA
TCTCCGACAAGAACAAGACCAACCCCGTGGAAATCCCCATCAAGCAGACAATCCCTAACTTCTT
CTTCGGCAGAGACACCGCCGAAGACTATGACGACCTGGACTACTGA

SEQ ID NO:16
>G5010 (synthetic protein) [Pan-Influenza Antigen]
MPMGSLQPLATLYLLGMLVASCLG
ASQGTKRSYEQMETGGERQNATEIRASVGRMVGGIGRFYIQMCTELKLSDYEGRLIQNSITIER
MVLSAFDERRNKYLEEHPSAGKDPKKTGGPIYRRRDGKWVRELILYDKEEIRRIWRQANNGEDA
TAGLTHLMIWHSNLNDATYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGVGTMVMEL
IRMIKRGINDRNFWRGENGRRTRIAYERMCNILKGKFQTAAQRAMMDQVRESRNPGNAEIEDLI
FLARSALILRGSVAHKSCLPACVYGLAVASGYDFEREGYSLVGIDPFRLLQNSQVFSLIRPNEN
PAHKSQLVWMACHSAAFEDLRVSSFIRGTRVVPRGQLSTRGVQIASNENMETMDSSTLELRSRY
WAIRTRSGGTTNQQRASAGQISVQPTFSVQRNLPFERATIMAAFTGNTEGRTSDMRTEIIRMME
SAKPEDVSFQGRGVFELSDEKATNPIVPSFDMSNEGSYFFGDNAEEYDN
MSNMDIDGINTGTIDKTPEEITPGTSGTTRPIIRPATLAPPSNKRTRNPSPERATTSSEDDVGR
KTQKKQTPTEIKKSVYNMVVKLGEFYNQMMVKAGLNDDMERNLIQNAHAVERILLAATDDKKTE
FQKKKNARDVKEGKEEIDHNKTGGTFYKMVRDDKTIYFSPIRITFLKEEVKTMYKTTMGSDGFS
GLNHIMIGHSQMNDVCFQRSKALKRVGLDPSLISTFAGSTIPRRSGATGVAIKGGGTLVAEAIR
FIGRAMADRGLLRDIKAKTAYEKILLNLKNKCSAPQQKALVDQVIGSRNPGIADIEDLTLLARS
MVVVRPSVASKVVLPISIYAKIPQLGFNVEEYSMVGYEAMALYNMATPVSILRMGDDAKDKSQL
FFMSCFGAAYEDLRVLSALTGTEFKPRSALKCKGFHVPAKEQVEGMGAALMSIKLQFWAPMTRS
GGNEVGGDGGSGQISCSPVFAVERPIALSKQAVRRMLSMNIEGRDADVKGNLLKMMNDSMAKKT
SGNAFIGKKMFQISDKNKTNPVEIPIKQTIPNFFFGRDTAEDYDDLDY

SEQ ID NO:17
>Nucleoprotein of Influenza A virus (H2N2) (Swiss-Protein ID
P21433) [JPT peptide Product Code: PM-INFA-NPH2N2]
ASQGTKRSYEQMETDGERQNANEIRASVGKMIGGIGRFYIQMCTELKLSDYEGRLIQNSL
TIERMVLSAFDERRNKYLEEHPSAGKDPKKTGGPIYKRVDGKWMRELVLYDKEEIRRIWR
QANNGDDATAGLTHMMIWHSNLNDTTYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGA
AVKGVGTMVMELIRMIKRGINDRNFWRGENGRKTRNAYERMCNILKGKFQTAAQRAMMDQ

VRESRNPGNAEIEDLIFLARSALILRGSVAHKSCLPACVYGPAVASGYDFEKEGYSLVGI
DPFKLLQNSQVYSLIRPNENPAHKSQLVWMACNSAAFEDLRVSSFIRGTKVIPRGKLSTR
GVQIASNENMDTMGSSTLELRSRYWAIRTRSGGNTNQQRASAGQISVQPTFSVQRNLPFD
KPTIMAAFTGNAEGRTSDMRAEIIRMMEGAKPEEVSFQGRGVFELSDEKATNPIVPSFDM
SNEGSYFFGDNAEEYDN

SEQ ID NO:18
>Nucleoprotein fragment of Zaire ebolavirus
DSRPQKIWMAPSLTESDMDYHKILTAGLSVQQGIVRQRVIPVYQVNNLEEICQLIIQAFEAGVD
FQESADSFLLMLCLHHAYQGDYKLFLESGAVKYLEGHGFRFEVKKRDGVKRLEELLPAVSSGKN
IKRTLAAMPEEETTEANAGQFLSFASLFLPKLVVGEKACLEKVQRQIQVHAEQGLIQYPTAWQS
VGHMMVIFRLMRTNFLIKFLLIHQGMHMVAGHDANDAVISNSVAQARFSGLLIVKTVLDHILQK
TERGVRLHPLARTAKVKNEVNSFKAALSSLAKHGEYAPFARLLNLSGVNNLEHGLFPQLSAIAL
GVATAHGSTLAGVNVGEQYQQLREAATEAEKQLQQYAESRELDHLGLDDQEKKILMNFHQKKNE
ISFQQTNAMVTLRKERLAKLTEAITAASLPKTSGHYDDDDDIPFPGPINDDDNPGHQDDDPTDS
QDTTIPDVVVDPDDGSYGEYQSYSENGMNAPDDLVLFDLDEDDEDTKPVPNRSTKGGQQKNSQK
GQHIEGRQTQFRPIQNVPGPHRTIHHASAPLTDNDRRNEPSGSTSPRMLTPINEEADPLDDADD
ETSSLPPLESDDEEQDRDGTSNRTPTVAPPAPVYRDHSEKKELPQDEQQDQDHTQEARNQDSDN
TQSEHSLEEMYRHILRSQGPFDAVLYYHMMKDEPVVFSTSDGKEYTYPDSLEEEYPPWLTEKEA
MNEENRFVTLDGQQFYWPVMNHKNKFMAILQHHQ

SEQ ID NO:19
>Nucleoprotein fragment of Sudan ebolavirus
AKLTEAITTASKIKVGDRYPDDNDIPFPGPIYDDTHPNPSDDNPDDSRDTTIPGGVVDPYDDES
NNYPDYEDSAEGTTGDLDLFNLDDDDDDSRPGPPDRGQNKERAARTYGLQDPTLDGAKKVPELT
PGSHQPGNLHITKSGSNTNQPQGNMSSTLHSMTPIQEESEPDDQKDNDDESLTSLDSEGDEDGE
SISEENTPTVAPPAPVYKDTGVDTNQQNGPSSTVDSQGSESEALPINSKKSSALEETYYHLLKT
QGPFEAINYYHLMSDEPIAFSTESGKEYIFPDSLEEAYPPWLSEKEALEKENRYLVIDGQQFLW
PVMSLRDKFLAVLQHD

SEQ ID NO:20
>Nucleoprotein fragment of Bundibugyo ebolavirus
AKLTEAITSTSILKTGRRYDDDNDIPFPGPINDNENSGQNDDDPTDSQDTTIPDVIIDPNDGGY
NNYSDYANDAASAPDDLVLFDLEDEDDADNPAQNTPEKNDRPATTKLRNGQDQDGNQGETASPR
VAPNQYRDKPMPQVQDRSENHDQTLQTQSRVLTPISEEADPSDHNDGDNESIPPLESDDEGSTD
TTAAETKPATAPPAPVYRSISVDDSVPSENIPAQSNQTNNEDNVRNNAQSEQSIAEMYQHILKT
QGPFDAILYYHMMKEEPIIFSTSDGKEYTYPDSLEDEYPPWLSEKEAMNEDNRFITMDGQQFYW
PVMNHRNKFMAILQHHR

SEQ ID NO:21
>Nucleoprotein fragment of Taï Forest ebolavirus
LVLFDLEDGDEDDHRPSSSSENNNKHSLTGTDSNKTSNWNRNPTNMPKKDSTQNNDNPAQRAQE
YARDNIQDTPTPHRALTPISEETGSNGHNEDDIDSIPPLESDEENNTETTITTTKNTTAPPAPV
YRSNSEKEPLPQEKSQKQPNQVSGSENTDNKPHSEQSVEEM

SEQ ID NO:22
>Pan-Ebola_Antigen (synthetic protein)

MPMGSLQPLATLYLLGMLVASCLGDSRPQKIWMAPSLTESDMDYHKILTAGLSVQQGIVRQRVI
PVYQVNNLEEICQLIIQAFEAGVDFQESADSFLLMLCLHHAYQGDYKLFLESGAVKYLEGHGFR
FEVKKRDGVKRLEELLPAVSSGKNIKRTLAAMPEEETTEANAGQFLSFASLFLPKLVVGEKACL
EKVQRQIQVHAEQGLIQYPTAWQSVGHMMVIFRLMRTNFLIKFLLIHQGMHMVAGHDANDAVIS
NSVAQARFSGLLIVKTVLDHILQKTERGVRLHPLARTAKVKNEVNSFKAALSSLAKHGEYAPFA
RLLNLSGVNNLEHGLFPQLSAIALGVATAHGSTLAGVNVGEQYQQLREAATEAEKQLQQYAESR
ELDHLGLDDQEKKILMNFHQKKNEISFQQTNAMVTLRKERLAKLTEAITAASLPKTSGHYDDDD
DIPFPGPINDDDNPGHQDDDPTDSQDTTIPDVVVDPDDGSYGEYQSYSENGMNAPDDLVLFDLD
EDDEDTKPVPNRSTKGGQQKNSQKGQHIEGRQTQFRPIQNVPGPHRTIHHASAPLTDNDRRNEP
SGSTSPRMLTPINEEADPLDDADDETSSLPPLESDDEEQDRDGTSNRTPTVAPPAPVYRDHSEK
KELPQDEQQDQDHTQEARNQDSDNTQSEHSLEEMYRHILRSQGPFDAVLYYHMMKDEPVVFSTS
DGKEYTYPDSLEEEYPPWLTEKEAMNEENRFVTLDGQQFYWPVMNHKNKFMAILQHHQAKLTEA
ITTASKIKVGDRYPDDNDIPFPGPIYDDTHPNPSDDNPDDSRDTTIPGGVVDPYDDESNNYPDY
EDSAEGTTGDLDLFNLDDDDDDSRPGPPDRGQNKERAARTYGLQDPTLDGAKKVPELTPGSHQP
GNLHITKSGSNTNQPQGNMSSTLHSMTPIQEESEPDDQKDNDDESLTSLDSEGDEDGESISEEN
TPTVAPPAPVYKDTGVDTNQQNGPSSTVDSQGSESEALPINSKKSSALEETYYHLLKTQGPFEA
INYYHLMSDEPIAFSTESGKEYIFPDSLEEAYPPWLSEKEALEKENRYLVIDGQQFLWPVMSLR
DKFLAVLQHDAKLTEAITSTSILKTGRRYDDDNDIPFPGPINDNENSGQNDDDPTDSQDTTIPD
VIIDPNDGGYNNYSDYANDAASAPDDLVLFDLEDEDDADNPAQNTPEKNDRPATTKLRNGQDQD
GNQGETASPRVAPNQYRDKPMPQVQDRSENHDQTLQTQSRVLTPISEEADPSDHNDGDNESIPP
LESDDEGSTDTTAAETKPATAPPAPVYRSISVDDSVPSENIPAQSNQTNNEDNVRNNAQSEQSI
AEMYQHILKTQGPFDAILYYHMMKEEPIIFSTSDGKEYTYPDSLEDEYPPWLSEKEAMNEDNRF
ITMDGQQFYWPVMNHRNKFMAILQHHRLVLFDLEDGDEDDHRPSSSSENNNKHSLTGTDSNKTS
NWNRNPTNMPKKDSTQNNDNPAQRAQEYARDNIQDTPTPHRALTPISEETGSNGHNEDDIDSIP
PLESDEENNTETTITTTKNTTAPPAPVYRSNSEKEPLPQEKSQKQPNQVSGSENTDNKPHSEQS
VEEM

SEQ ID NO:23
>Pan-Ebola_Antigen (synthetic mRNA)
ATGcccatggggtctctgcaaccgctggccaccttgtacctgctggggatgctggtcgcttcct
gcctcggagattctcgtcctcagaaaatctggatggcgccgagtctcactgaatctgacatgga
ttaccacaagatcttgacagcaggtctgtccgttcaacaggggattgttcggcaaagagtcatc
ccagtgtatcaagtaaacaatcttgaagaaatttgccaacttatcatacaggcctttgaagcag
gtgttgattttcaagagagtgcggacagtttccttctcatgctttgtcttcatcatgcgtacca
gggagattacaaacttttcttggaaagtggcgcagtcaagtatttggaagggcacgggttccgt
tttgaagtcaagaagcgtgatggagtgaagcgccttgaggaattgctgccagcagtatctagtg
gaaaaaacattaagagaacacttgctgccatgccggaagaggagacaactgaagctaatgccgg
tcagtttctctcctttgcaagtctattccttccgaaattggtagtaggagaaaaggcttgcctt
gagaaggttcaaaggcaaattcaagtacatgcagagcaaggactgatacaatatccaacagctt
ggcaatcagtaggacacatgatggtgattttccgtttgatgcgaacaaattttctgatcaaatt
tctcctaatacaccaagggatgcacatggttgccgggcatgatgccaacgatgctgtgatttca
aattcagtggctcaagctcgtttttcaggcttattgattgtcaaaacagtacttgatcatatcc
tacaaaagacagaacgaggagttcgtctccatcctcttgcaaggaccgccaaggtaaaaaatga
ggtgaactcctttaaggctgcactcagctccctggccaagcatggagagtatgctcctttcgcc
cgacttttgaacctttctggagtaaataatcttgagcatggtcttttccctcaactatcggcaa
ttgcactcggagtcgccacagcacacgggagtaccctcgcaggagtaaatgttggagaacagta
tcaacaactcagagaggctgccactgaggctgagaagcaactccaacaatatgcagagtctcgc

gaacttgaccatcttggacttgatgatcaggaaaagaaaattcttatgaacttccatcagaaaa
agaacgaaatcagcttccagcaaacaaacgctatggtaactctaagaaaagagcgcctggccaa
gctgacagaagctatcactgctgcgtcactgcccaaaacaagtggacattacgatgatgatgac
gacattccctttccaggacccatcaatgatgacgacaatcctggccatcaagatgatgatccga
ctgactcacaggatacgaccattcccgatgtggtggttgatcctgatgatggaagctacggcga
ataccagagttactcggaaaacggcatgaatgcaccagatgacttggtcctattcgatctagac
gaggacgacgaggacactaagccagtgcctaatagatcgaccaagggtggacaacagaagaaca
gtcaaaagggccagcatatagagggcagacagacacaattcaggccaattcaaaatgtcccagg
ccctcacagaacaatccaccacgccagtgcgccactcacggacaatgacagaagaaatgaaccc
tccggctcaaccagccctcgcatgctgacaccaattaacgaagaggcagacccactggacgatg
ccgacgacgagacgtctagccttccgcccttggagtcagatgatgaagagcaggacagggacgg
aacttccaaccgcacacccactgtcgccccaccggctcccgtatacagagatcactctgaaaag
aaagaactcccgcaagacgagcaacaagatcaggaccacactcaagaggccaggaaccaggaca
gtgacaacacccagtcagaacactcccttgaggagatgtatcgccacattctaagatcacaggg
gccatttgatgctgttttgtattatcatatgatgaaggatgagcctgtagttttcagtaccagt
gatggcaaagagtacacgtatccagactcccttgaagaggaatatccaccatggctcactgaaa
aagaggctatgaatgaagagaatagatttgttacattggatggtcaacaattttattggccggt
gatgaatcacaagaataaattcatggcaatcctgcaacatcatcaggctaaattgaccgaagcc
atcacgactgcatcgaagatcaaggttggagaccgttatcctgatgacaatgatattccatttc
ccgggccgatctatgatgacactcaccccaatccctctgatgacaatcctgatgattcacgtga
tacaactattccaggtggtgttgttgacccgtatgatgatgagagtaataattatcctgactac
gaggattcggctgaaggcaccacaggagatcttgatctcttcaatttggacgacgacgatgatg
acagccgaccaggaccaccagacaggggcagaacaaggagagggcggcccggacatatggcct
ccaagatccgaccttggacggagcgaaaaaggtgccggagttgaccccaggttcccatcaacca
ggcaacctccacatcaccaagtcgggttcaaacaccaaccaaccacaaggcaatatgtcatcta
ctctccatagtatgacccctatacaggaagaatcagagcccgatgatcaaaaagataatgatga
cgagagtctcacatcccttgactctgaaggtgacgaagatggtgagagcatctctgaggagaac
accccaactgtagctccaccagcaccagtctacaaagacactggagtagacactaatcagcaga
atggaccaagcagtactgtagatagtcaaggttctgaaagtgaagctctcccaatcaactctaa
aaagagttccgcactagaagaaacatattatcatctcctaaaaacacagggtccatttgaggca
atcaattattatcacctaatgagtgatgaacccattgcttttagcactgaaagtggcaaggaat
atatctttccagactcccttgaagaagcctacccgccgtggttgagtgagaaggaggccttaga
gaaggaaaatcgttatctggtcattgatggccagcaattcctctggccggtaatgagcctacgg
gacaagttccttgccgttcttcaacatgacgccaaattgaccgaagctattacttccacctcta
tcctcaaaacaggaaggcggtatgatgatgacaatgacataccctttccagggccaatcaatga
taacgagaactctggtcagaacgatgacgatccaacagactcccaggataccacaatcccggat
gtaataatcgatccaaacgatggtgggtataataattacagcgattatgcaaatgatgctgcaa
gtgctcctgatgacctagttctttttgaccttgaggacgaggatgatgctgataacccggctca
aaacacgccagaaaaaaatgatagaccagcaacaacaaagctgagaaatggacaggaccaggat
ggaaaccaaggcgaaactgcatccccacgggtagcccccaaccaatacagagacaagccaatgc
cacaagtacaggacagatccgaaaatcatgaccaaacccttcaaacacagtccagggttttgac
tcctatcagcgaggaagcagaccccagcgaccacaacgatggtgacaatgaaagcattcctccc
ctggaatcagacgacgagggtagcactgatactactgcagcagaaacaaagcctgccactgcac
ctcccgctcccgtctaccgaagtatctccgtagatgattctgtcccctcagagaacattcccgc
acagtccaatcaaacgaacaatgaggacaatgtcaggaacaatgctcagtcggagcaatccatt
gcagaaatgtatcaacatatcttgaaaacacaaggacctttgatgccatcctttactaccata
tgatgaaagaagagcccatcattttcagcactagtgatgggaaggagtatacatatccagactc

tcttgaagatgagtatccaccctggctcagcgagaaggaagccatgaacgaagacaatagattc
ataaccatggatggtcagcagttttactggcctgtgatgaatcatagaaataaattcatggcaa
tcctccagcatcacaggcttgttctttttgaccttgaagatggtgacgaggatgatcaccgacc
gtcaagttcatcagagaacaacaacaaacacagtcttacaggaactgacagtaacaaaacaagt
aactggaatcgaaacccgactaatatgccaaagaaagactccacacaaaacaatgacaatcctg
cacagcgggctcaagaatacgccagggataacatccaggatacaccaacaccccatcgagctct
aactcccatcagcgaagaaaccggctccaatggtcacaatgaagatgacattgatagcatccct
cctttggaatcagacgaagaaaacaacactgagacaaccattaccaccacaaaaaataccactg
ctccaccagcacctgtttatcggagtaattcagaaaaggagcccctcccgcaagaaaaatccca
gaagcaaccaaaccaagtgagtggtagtgagaataccgacaataaacctcactcagagcaatca
gtggaagaaatgTAA

SEQ ID NO:24
>G5003o (synthetic mRNA) [CD5sp+RBD2(Omicron)]
ATGcccatggggtctctgcaaccgctggccaccttgtacctgctggggatgctggtcgcttcct
gcctcggaagagtccaaccaacagaatctattgttagatttcctaatattacaaacttgtgccc
ttttgatgaagtttttaacgccaccagatttgcatctgtttatgcttggaacaggaagagaatc
agcaactgtgttgctgattattctgtcctatataatctcgcaccatttttcacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgt
aattagaggtgatgaagtcagacaaatcgctccagggcaaactggaaacattgctgattataat
tataaattaccagatgattttacaggctgcgttatagcttggaattctaacaagcttgattcta
aggttagtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttga
gagagatatttcaactgaaatctatcaggccggtaacaaaccttgtaatggtgttgcaggtttt
aattgttactttcctttacgatcatatagtttccgacccacttatggtgttggtcaccaaccat
acagagtagtagtactttcttttgaacttctacatgcaccagcaactgtttgtggacctaaaaa
gtctactaatttggttaaaaacaaatgtgtcaatttcTAA

SEQ ID No:25
>G5003o (synthetic protein) [CD5sp+RBD2 (Omicron)]
MPMGSLQPLATLYLLGMLVASCLGRVQPTESIVRFPNITNLCPFDEVFNATRFASVYAWNRKRI
SNCVADYSVLYNLAPFFTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGNIADYN
YKLPDDFTGCVIAWNSNKLDSKVSGNYNYLYRLFRKSNLKPFERDISTEIYQAGNKPCNGVAGF
NCYFPLRSYSFRPTYGVGHQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNF

SEQ ID No:26
>G5006d (synthetic mRNA)
atgcccatggggtctctgcaaccgctggccaccttgtacctgctggggatgctggtcgcttcct
gcctcggaagagtccaaccaacagaatctattgttagatttcctaatattacaaacttgtgccc
ttttggtgaagtttttaacgccaccagatttgcatctgtttatgcttggaacaggaagagaatc
agcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgt
aattagaggtgatgaagtcagacaaatcgctccagggcaaactggaaagattgctgattataat
tataaattaccagatgattttacaggctgcgttatagcttggaattctaacaatcttgattcta
aggttggtggtaattataattaccggtatagattgtttaggaagtctaatctcaaacctttga
gagagatatttcaactgaaatctatcaggccggtagcaaaccttgtaatggtgttgaaggtttt
aattgttactttcctttacaatcatatggtttccaacccactaatggtgttggttaccaaccat
acagagtagtagtactttcttttgaacttctacatgcaccagcaactgtttgtggacctaaaaa

gtctactaatttggttaaaaacaaatgtgtcaatttcTCAGGTGGTGGCGGTTCAGGCGGAGGT
GGCTCTGGCGGTGGCGGATCGatgtctgataatggaccccaaaatcagcgaaatgcaccccgca
ttacgtttggtggaccctcagattcaactggcagtaaccagaatggagaacgcagtggggcgcg
atcaaaacaacgtcggccccaaggtttacccaataatactgcgtcttggttcaccgctctcact
caacatggcaaggaaggccttaaattccctcgaggacaaggcgttccaattaacaccaatagca
gtccagatgaccaaattggctactaccgaagagctaccagacgaattcgtggtggtgacggtaa
aatgaaagatctcagtccaagatggtatttctactacctaggaactgggccagaagctggactt
ccctatggtgctaacaaagacggcatcatatgggttgcaactgagggagccttgaatacaccaa
aagatcacattggcacccgcaatcctgctaacaatgctgcaatcgtgctacaacttcctcaagg
aacaacattgccaaaaggcttctacgcagaagggagcagaggcggcagtcaagcctcttctcgt
tcctcatcacgtagtcgcaacagttcaagaaattcaactccaggcagcagtatgggaacttctc
ctgctagaatggctggcaatggctgtgatgctgctcttgctttgctgctgcttgacagattgaa
ccagcttgagagcaaaatgtctggtaaaggccaacaacaacaaggccaaactgtcactaagaaa
tctgctgctgaggcttctaagaagcctcggcaaaaacgtactgccactaaagcatacaatgtaa
cacaagctttcggcagacgtggtccagaacaaacccaaggaaattttggggaccaggaactaat
cagacaaggaactgattacaaacattggccgcaaattgcacaatttgcccccagcgcttcagcg
ttcttcggaatgtcgcgcattggcatggaagtcacaccttcgggaacgtggttgacctacacag
gtgccatcaaattggatgacaaagatccaaatttcaaagatcaagtcattttgctgaataagca
tattgacgcatacaaaacattcccaccaacagagcctaaaaaggacaaaaagaagaaggcttat
gaaactcaagccttaccgcagagacagaagaaacagcaaactgtgactcttcttcctgctgcag
atttggatgatttctccaaacaattgcaacaatccatgagcagtgctgactcaactcaggccAT
GGCATCCCCTGCTGCACCTCGTGCTGTTTCCTTTGCCGATAACAATGATATAACAAATACAAAC
CTATCTCGAGGTAGAGGACGTAATCCAAAACCACGAGCTGCACCAAATAACACTGTCTCTTGGT
ACACTGGGCTTACCCAACACGGGAAAGTCCCTCTTACCTTTCCACCTGGGCAGGGTGTACCTCT
TAATGCCAATTCTACCCCTGCGCAAAATGCTGGGTATTGGCGGAGACAGGACAGAAAAATTAAT
ACCGGGAATGGAATTAAGCAACTGGCTCCCAGGTGGTACTTCTACTACACTGGAACTGGACCCG
AAGCAGCACTCCCATTCCGGGCTGTTAAGGATGGCATCGTTTGGGTCCATGAAGATGGCGCCAC
TGATGCTCCTTCAACTTTTGGGACGCGGAACCCTAACAATGATTCAGCTATTGTTACACAATTC
GCGCCCGGTACTAAGCTTCCTAAAAACTTCCACATTGAGGGGACTGGAGGCAATAGTCAATCAT
CTTCAAGAGCCTCTAGCTTAAGCAGAAACTCTTCCAGATCTAGTTCACAAGGTTCAAGATCAGG
AAACTCTACCCGCGGCACTTCTCCAGGTCCATCTGGAATCGGAGCAGTAGGAGGTGATCTACTT
TACCTTGATCTTCTGAACAGACTACAAGCCCTTGAGTCTGGCAAAGTAAAGCAATCGCAGCCAA
AAGTAATCACTAAGAAAGATGCTGCTGCTGCTAAAAATAAGATGCGCCACAAGCGCACTTCCAC
CAAAAGTTTCAACATGGTGCAAGCTTTTGGTCTTCGCGGACCAGGAGACCTCCAGGGAAACTTT
GGTGATCTTCAATTGAATAAACTCGGCACTGAGGACCCACGTTGGCCCCAAATTGCTGAGCTTG
CTCCTACAGCCAGTGCTTTTATGGGTATGTCGCAATTTAAACTTACCCATCAGAACAATGATGA
TCATGGCAACCCTGTGTACTTCCTTCGGTACAGTGGAGCCATTAAACTTGACCCAAAGAATCCC
AACTACAATAAGTGGTTGGAGCTTCTTGAGCAAAATATTGATGCCTACAAAACCTTCCCTAAGA
AGGAAAAGAAACAAAAGGCACCAAAAGAAGAATCAACAGACCAAATGTCTGAACCTCCAAAGGA
GCAGCGTGTGCAAGGTAGCATCACTCAGCGCACTCGCACCCGTCCAAGTGTTCAGCCTGGTCCA
ATGATTGATGTTAACACTGATTCTGGTGGCGGTGGCTCGGGCGGAGGTGGGTCGGGTGGCGGCG
GATCAgaagcaaaaccttctggctcagttgtggaacaggctgaaggtgttgaatgtgatttttc
acctcttctgtctggcacacctcctcaggtttataatttcaagcgtttggtttttaccaattgc
aattataatcttaccaaattgctttcacttttttctgtgaatgattttacttgtagtcaaatat
ctccagcagcaattgctagcaactgttattcttcactgattttggattacttttcatacccact
tagtatgaaatccgatctcagtgttagttctgctggtccaatatcccagtttaattataaacag
tccttttctaatcccacatgtttgattttagcgactgttcctcataaccttactactattacta

agcctcttaagtacagctatattaacaagtgctctcgtcttctttctgatgatcgtactgaagt
acctcagttagtgaacgctaatcaatactcaccctgtgtatccattgtcccatccactgtgtgg
gaagacggtgattattataggaaacaactatctccacttgaaggtggtggctggcttgttgcta
gtggctcaactgttgccatgactgagcaattacagatgggctttggtattacagttcaatatgg
tacagacaccaatagtgtttgccccaagcttgaatttgctaatgacacaaaaattgcctctcaa
ttaggcaattgcgtggaatatTAG

SEQ ID No:27
>G5006d (synthetic protein)
MPMGSLQPLATLYLLGMLVASCLGRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRI
SNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYN
YKLPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLKPFERDISTEIYQAGSKPCNGVEGF
NCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFSGGGGSGGG
GSGGGGSMSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQGLPNNTASWFTALT
QHGKEGLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGL
PYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAEGSRGGSQASSR
SSSRSRNSSRNSTPGSSMGTSPARMAGNGCDAALALLLLDRLNQLESKMSGKGQQQQGQTVTKK
SAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASA
FFGMSRIGMEVTPSGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKAY
ETQALPQRQKKQQTVTLLPAADLDDFSKQLQQSMSSADSTQAMASPAAPRAVSFADNNDITNTN
LSRGRGRNPKPRAAPNNTVSWYTGLTQHGKVPLTFPPGQGVPLNANSTPAQNAGYWRRQDRKIN
TGNGIKQLAPRWYFYYTGTGPEAALPFRAVKDGIVWVHEDGATDAPSTFGTRNPNNDSAIVTQF
APGTKLPKNFHIEGTGGNSQSSSRASSLSRNSSRSSSQGSRSGNSTRGTSPGPSGIGAVGGDLL
YLDLLNRLQALESGKVKQSQPKVITKKDAAAAKNKMRHKRTSTKSFNMVQAFGLRGPGDLQGNF
GDLQLNKLGTEDPRWPQIAELAPTASAFMGMSQFKLTHQNNDDHGNPVYFLRYSGAIKLDPKNP
NYNKWLELLEQNIDAYKTFPKKEKKQKAPKEESTDQMSEPPKEQRVQGSITQRTRTRPSVQPGP
MIDVNTDSGGGGSGGGGSGGGGSEAKPSGSVVEQAEGVECDFSPLLSGTPPQVYNFKRLVFTNC
NYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSYPLSMKSDLSVSSAGPISQFNYKQ
SFSNPTCLILATVPHNLTTITKPLKYSYINKCSRLLSDDRTEVPQLVNANQYSPCVSIVPSTVW
EDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYGTDTNSVCPKLEFANDTKIASQ
LGNCVEY

SEQ ID No:28
>G50012 (synthetic mRNA)
atgcctatgggcagcctgcagccactggctacactgtacctgctgggcatgctggtggcctctt
gtctgggcataggagagtgtccaaagtatgtcaggagtgcaaaattaaggatggttacaggact
aaggaacatcccatccattcaatccagaggtttgtttggagccattgccggtttcattgaaggg
gggtggactggaatggtagatgggtggtatggttatcatcatcagaatgagcaaggatctggct
atgctgcagatcaaaaaagtacacaaaatgccattaacgggattacaaacaaggtgaattctgt
aattgagaaaatgaacactcaattcacagctgtgggcaaagaattcaacaaattggaaagaagg
atggaaaacttaaataaaaaagttgatgatgggtttctagacatttggacatataatgcagaat
tgttggttctactggaaaatgaaaggactttggatttccatgactccgccagccaaggcactaa
gagaagctacgagcagatggaaaccggaggcgaacggcagaacgccacagagatcagagcctct
gtgggccgtatggtcggcggcatcggcagattctacatccagatgtgcaccgaactgaagctga
gcgactacgagggccgcctgatccagaacagcatcacaatcgagagaatggtgctgtccgcctt
tgacgagcggagaaacaaatacctggaagagcaccctagcgccggaaaagatcctaagaaaacc
ggcggacctatctacagaagaagagatggtaagtgggtgagagagctgattctgtacgataagg

aagagattcgaagaatctggagacaggccaacaacggcgaggatgccaccgcaggcctgacaca
cctgatgatctggcacagcaacctgaacgatgcgacctaccagcgcacgcgggccctggtcaga
accggcatggatcctcggatgtgtagcctgatgcagggcagcacactgccaagacggagtgggg
ccgccggcgctgcagtgaagggcgtcggaaccatggtgatggagctgatccggatgataaagcg
gggcatcaacgacagaaacttctggcgaggcgagaacggccgaagaacccggatcgcctacgag
agaatgtgcaacatcctgaaaggaaaattccagaccgccgcccagcgggccatgatggaccagg
tgcgcgagagcagaaaccccggcaatgccgagatcgaggacctgatcttcctggccagaagcgc
cctcattcttagaggctctgtggcccacaagagctgtctgcctgcctgtgtgtacggcctggca
gtggcctcaggctacgacttcgagcgggaaggatacagtctggtgggcatcgacccctttcagac
tcctgcagaatagccaggtgtttagcctgatcagaccaaacgaaaaccccgcccataagagcca
gctggtgtggatggcctgccacagcgccgcctttgaggatctgagagtgagctcttttatcaga
ggcacccgggtggttccacgaggtcaactgtctacaagaggtgtgcagatcgccagcaacgaga
acatggagaccatggatagcagcaccctggaactgagatccagatactgggccatcaggacacg
gagcggcggcaccaccaatcagcagcgcgccagcgccggccagatctctgtccagcctacgttt
agcgtgcagcggaatttgcccttcgaacgcgccacaatcatggctgctttcaccggcaatacag
agggcagaaccagcgatatgagaacagaaattatccgtatgatggagtccgcaaaacctgagga
cgtgtccttccaaggcagaggcgtgttcgagctgagcgacgagaaggccaccaaccctatcgtg
cctagcttcgatatgtctaatgagggcagctacttttttcggagataacgccgaagagtacgaca
acatgtctaatatggatatcgacggcattaacaccggcaccatcgacaaaaccccctgaggagat
cacccctggcaccagcggcacaacccggcccatcatccgccccgctacactggctccacctagc
aacaagcggaccagaaatccctcgccagaaagagccacaacctccagcgaggacgacgtgggac
ggaagacacaaaagaagcagacccctacagagatcaagaagtctgtttacaacatggtggtgaa
actgggcgagttctacaaccagatgatggtgaaggccggcctgaacgacgatatggaaagaaat
ctgatccagaacgcccacgccgtggagcggattctgctggccgccaccgatgataagaagaccg
aattccagaaaaagaaaaacgccagagacgtgaaggaaggcaaggaagagatcgaccacaacaa
gacaggcggcacattctacaagatggtccgggacgacaagaccatctacttcagccctatccgg
ataacattcctgaaagaagaagtgaagaccatgtacaaaaccacaatgggctctgacggcttca
gcggcctgaatcacatcatgatcggccactctcaaatgaacgatgtgtgcttccagagaagcaa
ggctctgaagcgcgtgggcctggatcctagcctgatctctaccttcgccggcagcaccatcccc
agaagatcgggcgctaccggcgtggctatcaagggaggaggcacactggtggctgaagccatca
gattcatcggaagagccatggccgacagaggactcctgagagatatcaaagccaaaaccgccta
cgaaaaaatcctgctgaacctgaagaacaagtgcagcgcgcctcaacagaaggccctggtggac
caggttatcggctctagaaaccctggaatcgccgatatcgaggacctgacactgctggccagat
ctatggtggtggtgagaccctccgtggccagcaaggtggtgctgcctatcagcatctacgccaa
gatccctcagctgggatttaacgtggaagaatacagcatggttggttatgaggccatggccctg
tacaacatggccacacctgtgtccatcctgagaatgggcgacgatgccaaagacaagagccagc
tgttcttcatgagctgcttcggcgctgcctatgaggacctgagagtgctgtccgctcttacagg
aacagagttcaagcctaggagcgcactgaagtgcaagggcttccacgtgcccgccaaggaacag
gtggaaggcatgggagctgctctgatgtccatcaagctgcaattttgggctcctatgacccgga
gcggcggaaatgaggtgggtggcgacggaggcagcggacagatttcttgcagccccgtatttgc
cgtggagagaccaatcgccctgtccaagcaggccgtgagaagaatgctgagcatgaacatcgag
ggccgggacgccgacgtgaagggcaacctgttgaagatgatgaacgacagcatggccaagaaga
ccagtggcaatgccttcatcggcaagaagatgttccagatctccgacaagaacaagaccaaccc
cgtggaaatccccatcaagcagacaatccctaacttcttcttcggcagagacaccgccgaagac
tatgacgacctggactacataggaaattgcccaatatgggtgaaaacacctttgaagcttgcca
atggaaccaaatatagacctcctgcaaaactattaaaggaaaggggtttcttcggagctattgc
tggtttcctagaaggaggatgggaaggaatgattgcaggctggcacggatacacatctcacgga

gcacatggagtggcagtggcggcggaccttaagagtacgcaagaagctataaacaagataacaa
aaaatctcaattctttgagtgagctagaagtaaagaatcttcaaagactaagtggtgccatgga
tgaactccacaacgaaatactcgagctggatgagaaagtggatgatctcagagctgacactata
agctcgcaaatagaacttgcagtcttgctttccaacgaaggaataataaacagtgaagatgagt
ga

SEQ ID No:29
>G50012 (synthetic protein)
MPMGSLQPLATLYLLGMLVASCLGIGECPKYVRSAKLRMVTGLRNIPSIQSRGLFGAIAGFIEG
GWTGMVDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKLERR
MENLNKKVDDGFLDIWTYNAELLVLLENERTLDFHDSASQGTKRSYEQMETGGERQNATEIRAS
VGRMVGGIGRFYIQMCTELKLSDYEGRLIQNSITIERMVLSAFDERRNKYLEEHPSAGKDPKKT
GGPIYRRRDGKWVRELILYDKEEIRRIWRQANNGEDATAGLTHLMIWHSNLNDATYQRTRALVR
TGMDPRMCSLMQGSTLPRRSGAAGAAVKGVGTMVMELIRMIKRGINDRNFWRGENGRRTRIAYE
RMCNILKGKFQTAAQRAMMDQVRESRNPGNAEIEDLIFLARSALILRGSVAHKSCLPACVYGLA
VASGYDFEREGYSLVGIDPFRLLQNSQVFSLIRPNENPAHKSQLVWMACHSAAFEDLRVSSFIR
GTRVVPRGQLSTRGVQIASNENMETMDSSTLELRSRYWAIRTRSGGTTNQQRASAGQISVQPTF
SVQRNLPFERATIMAAFTGNTEGRTSDMRTEIIRMMESAKPEDVSFQGRGVFELSDEKATNPIV
PSFDMSNEGSYFFGDNAEEYDNMSNMDIDGINTGTIDKTPEEITPGTSGTTRPIIRPATLAPPS
NKRTRNPSPERATTSSEDDVGRKTQKKQTPTEIKKSVYNMVVKLGEFYNQMMVKAGLNDDMERN
LIQNAHAVERILLAATDDKKTEFQKKKNARDVKEGKEEIDHNKTGGTFYKMVRDDKTIYFSPIR
ITFLKEEVKTMYKTTMGSDGFSGLNHIMIGHSQMNDVCFQRSKALKRVGLDPSLISTFAGSTIP
RRSGATGVAIKGGGTLVAEAIRFIGRAMADRGLLRDIKAKTAYEKILLNLKNKCSAPQQKALVD
QVIGSRNPGIADIEDLTLLARSMVVVRPSVASKVVLPISIYAKIPQLGFNVEEYSMVGYEAMAL
YNMATPVSILRMGDDAKDKSQLFFMSCFGAAYEDLRVLSALTGTEFKPRSALKCKGFHVPAKEQ
VEGMGAALMSIKLQFWAPMTRSGGNEVGGDGGSGQISCSPVFAVERPIALSKQAVRRMLSMNIE
GRDADVKGNLLKMMNDSMAKKTSGNAFIGKKMFQISDKNKTNPVEIPIKQTIPNFFFGRDTAED
YDDLDYIGNCPIWVKTPLKLANGTKYRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHG
AHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTI
SSQIELAVLLSNEGIINSEDE

CLAIMS

What is claimed:

1. A composition for stimulating an immune response against a coronavirus in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':
(i) a nucleotide sequence encoding a mammalian signal peptide; and
(ii) a nucleotide sequence encoding a coronavirus nucleocapsid protein.

2. The composition of claim 1, wherein the coronavirus is a betacoronavirus, optionally wherein the betacoronavirus is a human betacoronavirus.

3. The composition of claim 2, wherein the betacoronavirus comprises a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), a middle east respiratory syndrome-related coronavirus (MERS-CoV), or a combination thereof.

4. The composition of claim 3, wherein the betacoronavirus comprises a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2).

5. The composition of claim 4, wherein the coronavirus nucleocapsid protein comprises a first nucleocapsid protein and a second nucleocapsid protein, wherein the first nucleocapsid protein is a SARS-CoV-2 nucleocapsid protein of a first variant from a first clade, and the second nucleocapsid protein is a SARS-CoV-2 nucleocapsid protein of a second variant from a second clade, and wherein the first clade and the second clade are different clades as defined by one or more of the World Health Organization, Pango, GISAID, and Nextstrain.

6. A composition for stimulating an immune response against a coronavirus in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':
(i) a nucleotide sequence encoding a mammalian signal peptide; and
(ii) a nucleotide sequence encoding two or more coronavirus nucleocapsid proteins.

7. The composition of claim 6, wherein the coronavirus is a betacoronavirus, optionally wherein the betacoronavirus is a human betacoronavirus.

8. The composition of claim 7, wherein the betacoronavirus comprises a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), a middle east respiratory syndrome-related coronavirus (MERS-CoV), or a combination thereof.

9. The composition of claim 8, wherein the betacoronavirus comprises a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2).

10. The composition of claim 9, wherein the two or more coronavirus nucleocapsid proteins comprise a SARS-CoV-2 nucleocapsid protein and a MERS nucleocapsid protein.

11. The composition of claim 9, wherein the two or more coronavirus nucleocapsid proteins comprise a SARS-CoV-2 nucleocapsid protein, a SARS-CoV-1 nucleocapsid protein, and a MERS nucleocapsid protein.

12. The composition of any one of claims 6-11, wherein the two or more coronavirus nucleocapsid proteins are separated by a linker of from one to ten residues in length.

13. The composition of any one of claims 1-12, wherein the mammalian signal peptide is a signal peptide of a surface protein expressed in mammalian antigen presenting cells.

14. The composition of claim 13, wherein the mammalian signal peptide is a CD5 signal peptide and the amino acid sequence of the CD5 signal peptide comprises SEQ ID NO:8, or the amino acid sequence at least 90% or 95% identical to SEQ ID NO:8.

15. The composition of any one of claims 1-14, wherein the amino acid sequence of the nucleocapsid protein comprises residues 2-419 of SEQ ID NO:5, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to residues 2-419 of SEQ ID NO:5.

16. The composition of any one of claims 1-14, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:6, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:6.

17. The composition of any one of claims 6-14, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:7, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:7.

18. The composition of claim 16, wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO:2.

19. The composition of claim 17, wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO:3 or SEQ ID NO:4.

20. The composition of any one of claims 1-14, wherein the amino acid sequence of the fusion protein comprises residues 2-413 of SEQ ID NO:9, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to residues 2-413 of SEQ ID NO:9.

21. The composition of any one of claims 1-14, wherein the amino acid sequence of the fusion protein comprises residues 2-422 of SEQ ID NO:10, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to residues 2-422 of SEQ ID NO:10.

22. The composition of any one of claims 1-21, wherein the composition does not comprise liposomes or lipid nanoparticles.

23. The composition of any one of claims 1-22, wherein the mRNA is a self-replicating mRNA.

24. The composition of claim 23, wherein the self-replicating RNA comprises an Alphavirus replicon lacking a viral structural protein coding region.

25. The composition of claim 24, wherein the Alphavirus is selected from the group consisting of a Venezuelan equine encephalitis virus, a Sindbis virus, and a Semliki Forrest virus.

26. The composition of claim 25, wherein the Alphavirus is a Venezuelan equine encephalitis virus.

27. The composition of any one of claims 23-26, wherein the Alphavirus replicon comprises a nonstructural protein coding region with an insertion of 12-18 nucleotides resulting in expression of a nonstructural Protein 2 (nsP2) comprising from 4 to 6 additional amino acids between beta sheet 4 and beta sheet 6 of the nsP2.

28. The composition of any one of claims 1-27, wherein the self-replicating mRNA is a temperature-sensitive agent (ts-agent) that is capable of expressing the fusion at a permissive temperature but not at a non-permissive temperature.

29. The composition of claim 28, wherein the permissive temperature is from 31°C to 35°C and the non-permissive temperature is at least 37°C ± 0.5°C.

30. A method for stimulating an immune response against a coronavirus in a mammalian subject, comprising administering the composition of any one of claims 1-29 to a mammalian subject so as to stimulate an immune response against the coronavirus nucleocapsid protein in the mammalian subject

31. The method of claim 30, wherein the composition is administered intradermally.

32. The method of claim 30 or claim 31, wherein the immune response comprises a coronavirus-reactive cellular immune response.

33. The method of claim 32, wherein the immune response further comprises a coronavirus-reactive humoral immune response.

34. The method of any one of claims 30-33, wherein the mammalian subject is a human subject.

35. A kit comprising:
the composition of any one of claims 1-29 or any one of claims 37-62; and
a device for intradermal delivery of the composition to a mammalian subject.

36. The kit of claim 35, wherein the device comprises a syringe and a needle.

37. A composition for stimulating an immune response against two or more viruses in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':
(i) a nucleotide sequence encoding a mammalian signal peptide; and
(ii) a nucleotide sequence encoding a first nucleocapsid protein of a first virus and a second nucleocapsid protein of a second virus.

38. The composition of claim 37, wherein the first and second viruses are capable of causing disease upon infection of a human subject.

39. The composition of claim 38, wherein the first and second viruses are different variants, subtypes or lineages of the same species.

40. The composition of claim 38, wherein the first and second viruses are different species of the same genus.

41. The composition of claim 40, wherein the first and second viruses are both members of the betacoronavirus genus.

42. The composition of claim 41, wherein the first and second viruses comprise a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) and a middle east respiratory syndrome-related coronavirus (MERS-CoV).

43. The composition of claim 38, wherein the first and second viruses are members of different families, orders, classes, or phyla of the same kingdom.

44. The composition of claim 43, wherein the first and second viruses are both members of the orthomyxoviridae family.

45. The composition of claim 44, wherein the first and second viruses comprise an influenza A virus and an influenza B virus.

46. The composition of claim 45, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:16, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:16.

47. The composition of claim 38, wherein the first and second viruses are both members of the orthornavirae kingdom, optionally wherein the first and second viruses comprise: (a) a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), or a middle east respiratory syndrome-related coronavirus (MERS-CoV); and (b) an influenza A virus or an influenza B virus.

48. The composition of claim 40, wherein the first and second viruses are both members of the ebolavirus genus, optionally wherein the first and second viruses are selected from the group consisting of Zaire ebolavirus, Sudan ebolavirus, Bundibugyo ebolavirus, and Taï Forest ebolavirus.

49. The composition of claim 48, wherein the nucleotide sequence further encodes a third nucleocapsid protein of a third virus and a fourth nucleocapsid protein of a fourth virus, and the

first, second, third and fourth viruses are Zaire ebolavirus, Sudan ebolavirus, Bundibugyo ebolavirus, and Taï Forest ebolavirus.

50. The composition of claim 49, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:22, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:22.

51. The composition of claim 49, wherein the nucleotide sequence (ii) encodes a shared portion of the first nucleocapsid protein of the first virus for stimulating an immune response against all of the first, second, third and fourth viruses.

52. The composition of claim 51, wherein the nucleotide sequence (ii) encodes an individual portion of each of the first, second, third and fourth nucleocapsid proteins for stimulating an immune response against all of the first, second, third and fourth viruses.

53. The composition of claim 52, wherein the nucleotide sequence (ii) encodes a fragment of the individual portion of the second nucleocapsid protein of the second virus for stimulating an immune response against the second and third viruses.

54. The composition of claim 37, wherein the nucleotide sequence (ii) encodes a shared portion of the first nucleocapsid protein of the first virus for stimulating an immune response against both the first and second viruses.

55. The composition of claim 54, wherein the nucleotide sequence (ii) encodes an individual portion of each of the first and second nucleocapsid proteins for stimulating an immune response against both the first and second viruses.

56. The composition of any one of claims 37-48, wherein the nucleotide sequence of (ii) further encodes at least one further nucleocapsid protein of at least one further virus, and wherein the at least one further virus is different from the first and second viruses.

57. The composition of any one of claims 37-56, wherein the first and second, or the first, second, and further nucleocapsid proteins are separated by a linker of from one to ten residues in length.

58. The composition of any one of claims 37-57, wherein the mammalian signal peptide is a signal peptide of a surface protein expressed in mammalian antigen presenting cells.

59. The composition of any one of claims 37-58, wherein the mRNA is a self-replicating mRNA.

60. The composition of claim 59, wherein the self-replicating mRNA is a temperature-sensitive agent (ts-agent) that is capable of expressing the fusion protein a permissive temperature but not at a non-permissive temperature.

61. The composition of claim 60, wherein the permissive temperature is from 31°C to 35°C and the non-permissive temperature is at least 37°C ± 0.5°C.

62. The composition of any one of claims 1-29 or any one of claims 37-61, wherein the composition further comprises chitosan.

63. A method for stimulating an immune response against two or more viruses in a mammalian subject, comprising administering the composition of any one of claims 37-62 to a mammalian subject to stimulate an immune response against the nucleocapsid proteins of the two or more viruses in the mammalian subject

64. The method of claim 63, wherein the composition is administered intradermally.

65. The method of claim 63 or claim 64, wherein the immune response comprises a cellular immune response reactive with the two or more viruses.

66. The method of claim 65, wherein the cellular immune response comprises a nucleocapsid protein-specific helper T lymphocyte (Th) response comprising nucleocapsid protein-specific cytokine secretion.

67. The method of claim 66, wherein nucleocapsid protein-specific cytokine secretion comprises secretion of one or both of interferon-gamma and interleukin-4.

68. The method of claim 65, wherein the cellular immune response comprises a nucleocapsid protein-specific cytotoxic T lymphocyte (CTL) response.

69. The method of any one of claims 65-68, wherein the immune response further comprises a humoral immune response reactive with the two or more viruses.

70. The method of any one of claims 63-69, wherein the mammalian subject is a human subject.

71. A composition for stimulating an immune response against a virus in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':

(i) a nucleotide sequence encoding a mammalian signal peptide;

(ii) a nucleotide sequence encoding a first viral antigen or fragment thereof of a first virus; and

(iii) a nucleotide sequence encoding a second viral antigen or fragment thereof of the first virus or a second virus,

wherein the first viral antigen is a nucleocapsid protein and the second viral antigen is a surface protein, or the first viral antigen is a surface protein and the second viral antigen is a nucleocapsid protein.

72. A composition for stimulating an immune response against two or more viruses in a mammalian subject, comprising an excipient, and a messenger RNA (mRNA) comprising an open reading frame (ORF) encoding a fusion protein, wherein the ORF comprises from 5' to 3':

(i) a nucleotide sequence encoding a mammalian signal peptide;

(ii) a nucleotide sequence encoding a first viral antigen or fragment thereof of a first virus;

(iii) a nucleotide sequence encoding a second viral antigen or fragment thereof of the first virus;

(iv) a nucleotide sequence encoding a third viral antigen or fragment thereof of a second virus;

(iii) a nucleotide sequence encoding a fourth viral antigen or fragment thereof of the second virus,

wherein the first viral antigen is a first nucleocapsid protein and the second viral antigen is a first surface protein, or the first viral antigen is a first surface protein and the second viral antigen is a first nucleocapsid protein, and

wherein the third viral antigen is a second nucleocapsid protein and the fourth viral antigen is a second surface protein, or the third viral antigen is a second surface protein and the fourth viral antigen is a second nucleocapsid protein.

73. The composition of claim 71 or claim 72, wherein the mRNA is a self-replicating mRNA.

74. The composition of claim 73, wherein the self-replicating RNA comprises an Alphavirus replicon lacking a viral structural protein coding region.

75. The composition of claim 74, wherein the Alphavirus is selected from the group consisting of a Venezuelan equine encephalitis virus, a Sindbis virus, and a Semliki Forrest virus.

76. The composition of claim 74, wherein the Alphavirus is a Venezuelan equine encephalitis virus.

77. The composition of any one of claims 73-76, wherein the self-replicating mRNA is a temperature-sensitive agent (ts-agent) that is capable of expressing the fusion protein at a permissive temperature but not at a non-permissive temperature.

78. The composition of claim 77, wherein the permissive temperature is from 31°C to 35°C, and the non-permissive temperature is at least 37°C ± 0.5°C.

79. The composition of any one of claims 74-78, wherein the Alphavirus replicon comprises a nonstructural protein coding region with an insertion of 12-18 nucleotides resulting in expression of a nonstructural Protein 2 (nsP2) comprising from 4 to 6 additional amino acids between beta sheet 4 and beta sheet 6 of the nsP2.

80. The composition of any one of claims 71-79, wherein the first virus and/or the second virus is a coronavirus, optionally wherein the coronavirus is a betacoronavirus, optionally wherein the betacoronavirus is a human betacoronavirus.

81. The composition of claim 80, wherein the first and/or the second virus is a betacoronavirus independently selected from the group consisting of a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), and a middle east respiratory syndrome-related coronavirus (MERS-CoV).

82. The composition of claim 80, wherein the first virus is SARS-CoV-2 and the second virus is MERS-CoV.

83. The composition of any one of claims 80-82, wherein the surface protein, the first surface protein and/or the second surface protein each comprise a receptor-binding domain (RBD) of a coronavirus Spike protein.

84. The composition of claim 83, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:27, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:27.

85. The composition of any one of claims 71-79, wherein the first virus and/or the second virus is a member of the orthomyxoviridae family.

86. The composition of claim 85, wherein the first and/or the second virus is independently selected from the group consisting of an influenza A virus (IAV) and an influenza B virus (IBV).

87. The composition of claim 86, wherein the first virus is IAV and the second virus is IBV.

88. The composition of any one of claims 85-87, wherein the surface protein, the first surface protein and/or the second surface protein each comprise a portion of an influenza hemagglutinin.

89. The composition of claim 88, wherein the amino acid sequence of the fusion protein comprises SEQ ID NO:29, or the amino acid sequence at least 75%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:29.

90. The composition of any one of claims 71-89, wherein the composition further comprises chitosan.

91. A kit comprising:
(i) the composition of any one of claims 71- 90; and
(ii) a device for intradermal delivery of the composition to a mammalian subject.

92. The kit of claim 91, wherein the device comprises a syringe and a needle.

93. The kit of claim 91 or claim 92, further comprising instructions for use of the device to administer the composition to a mammalian subject to stimulate an immune response against one or more of the first viral antigen, the second viral antigen, the third viral antigen, and the fourth viral antigen.

94. A method of stimulating an immune response in a mammalian subject, comprising administering the composition of any one of claims 71-90 to a mammalian subject to stimulate an immune response against one or more of the first viral antigen, the second viral antigen, the third viral antigen, and the fourth viral antigen in the mammalian subject.

95. The method of claim 94, wherein the composition is administered intradermally.

96. The method of claim 95, wherein the immune response comprises a cellular immune response reactive against one or more of the first viral antigen, the second viral antigen, the third viral antigen, and the fourth viral antigen.

97. The method of claim 96, wherein the immune response further comprises a humoral immune response reactive against one or more of the first viral antigen, the second viral antigen, the third viral antigen, and the fourth viral antigen.

98. The method of any one of claims 94-97, wherein the mammalian subject is a human subject.

99. A method for active booster immunization against at least one virus, comprising intradermally administering the composition of any one of claims 1-29, any one of claims 37-62, or any one of claims 71-90 to a mammalian subject in need thereof to stimulate a secondary immune response against the virus, wherein the mammalian subject had already undergone a primary immunization regimen against the virus.

100. The method of claim 99, wherein the primary immunization regimen comprises administration of at least one dose of a different vaccine against the virus.

101. The method of claim 100, wherein the different vaccine comprises a protein antigen of the at least one virus, optionally wherein the protein antigen is a recombinant protein or fragment thereof, or an inactivated virus.

102. A method for active booster immunization against at least one virus, comprising:

(i) intradermally administering the composition of any one of claims 1-29, any one of claims 37-62, or any one of claims 71-90 to a mammalian subject in need thereof to stimulate a primary immune response against the virus; and

(ii) administering at least one dose of a different vaccine against the virus to the mammalian subject to stimulate a secondary immune response against the virus.

103. The method of claim 102, wherein the different vaccine comprises a protein antigen of the at least one virus, optionally wherein the protein antigen is a recombinant protein or fragment thereof, or an inactivated virus.

104. A method for active primary immunization against at least one virus, comprising:

(i) intradermally administering the composition of any one of claims 1-29, any one of claims

37-62, or any one of claims 71-90 to a mammalian subject in need thereof to stimulate a primary immune response against the virus; wherein the mammalian subject had not undergone a primary immunization regimen against the virus.

105. The method of claim 104, further comprising:

(ii) administering at least one dose of a different vaccine against the virus to the mammalian subject to stimulate a secondary immune response against the virus.

106. The method of claim 105, wherein the different vaccine comprises a protein antigen of the at least one virus, optionally wherein the protein antigen is a recombinant protein or fragment thereof, or an inactivated virus.

107. The method of any one of claims 94-106, wherein the mammalian subject is a human subject.

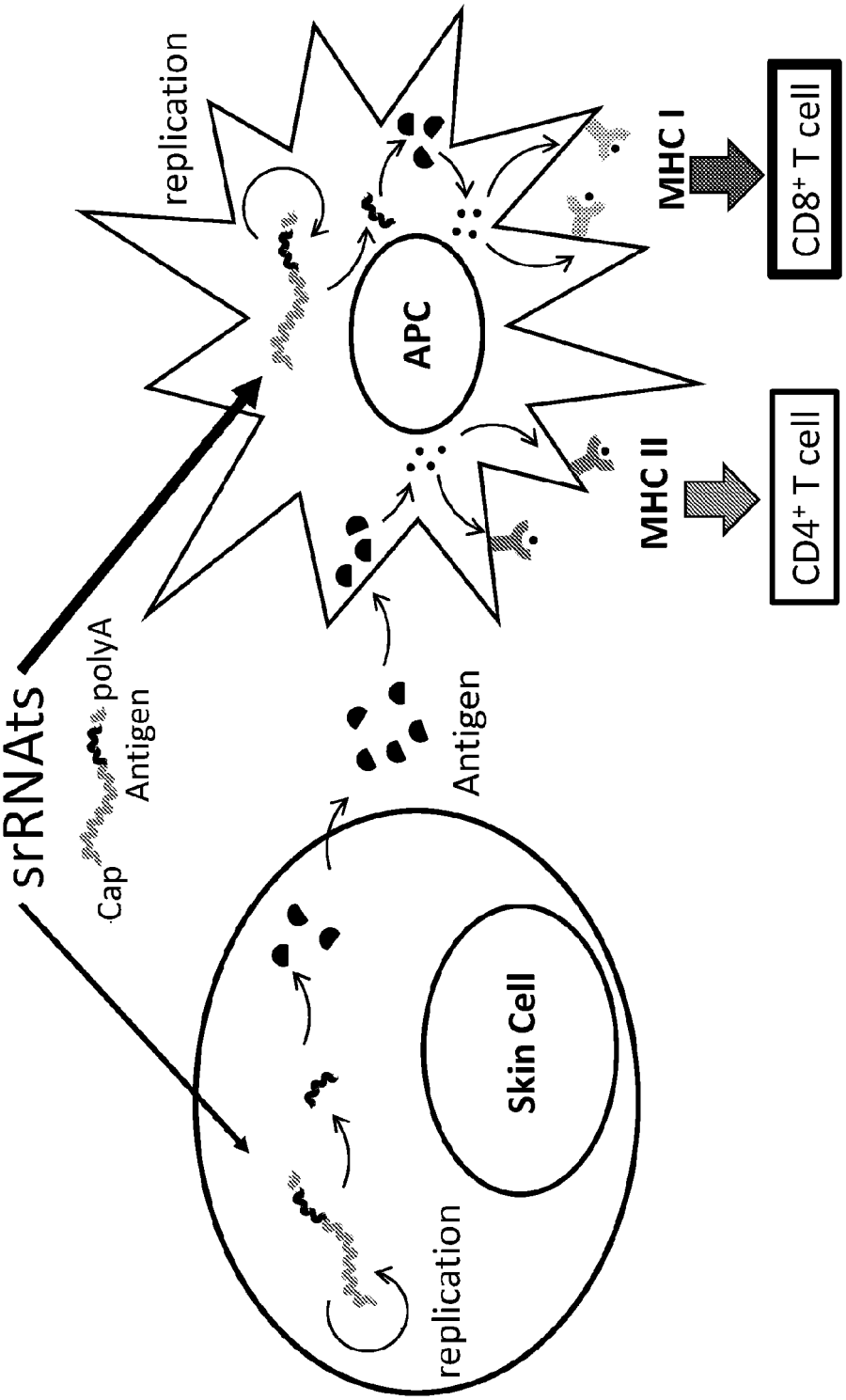

srRNAts
Cap
polyA
Antigen
replication
Skin Cell
Antigen
APC
MHC I
MHC II
CD8+ T cell
CD4+ T cell

FIG. 1

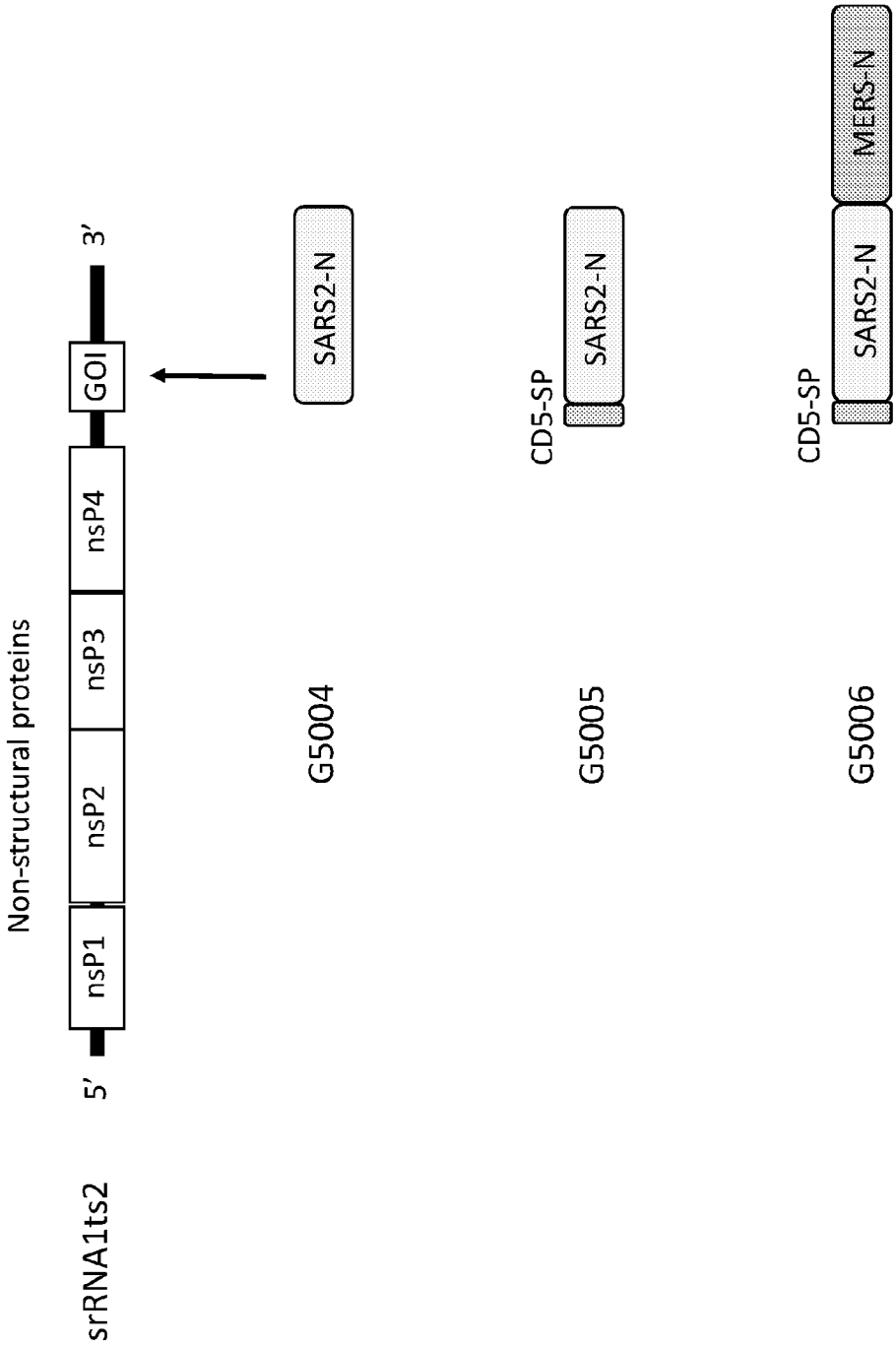

srRNA1ts2
5'
nsP1
nsP2
nsP3
nsP4
Non-structural proteins
GOI
3'
SARS2-N
G5004
CD5-SP
SARS2-N
G5005
CD5-SP
SARS2-N
MERS-N
G5006

FIG. 2

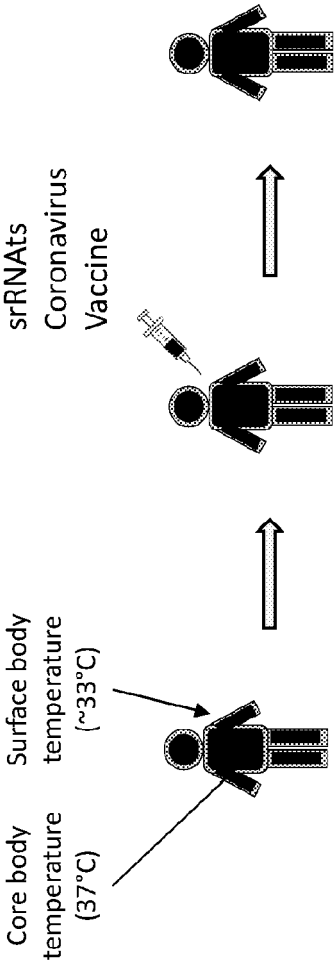

Core body temperature (37°C)
Surface body temperature (~33°C)
srRNAts Coronavirus Vaccine

FIG. 3

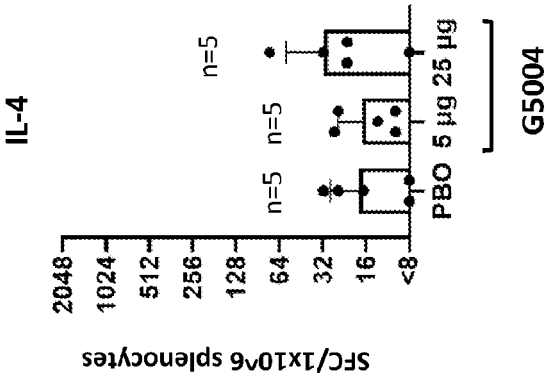

IL-4
SFC/1x10^6 splenocytes
2048
1024
512
256
128
64
32
16
<8
n=5
n=5
n=5
PBO
5 µg
25 µg
G5004

FIG. 4B

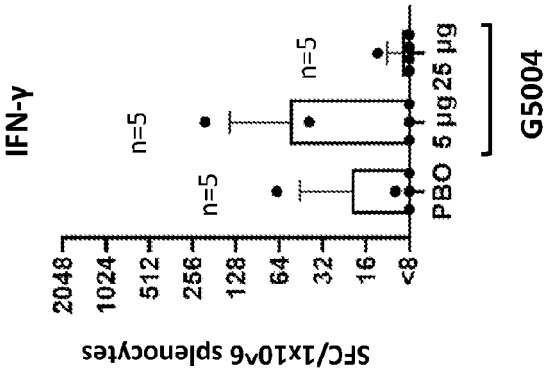

IFN-γ
SFC/1x10^6 splenocytes
2048
1024
512
256
128
64
32
16
<8
n=5
n=5
n=5
PBO
5 µg
25 µg
G5004

FIG. 4A

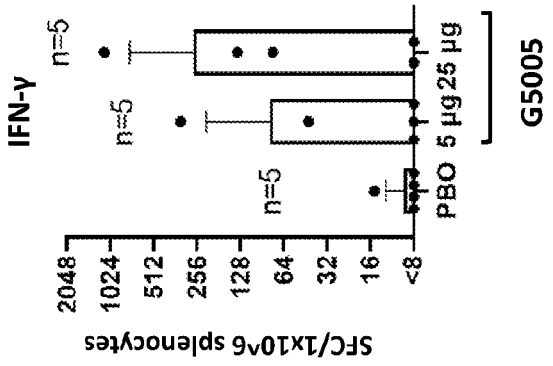

IFN-γ
SFC/1x10^6 splenocytes
2048
1024
512
256
128
64
32
16
<8
n=5
n=5
n=5
PBO
5 µg
25 µg
G5005

FIG. 5A

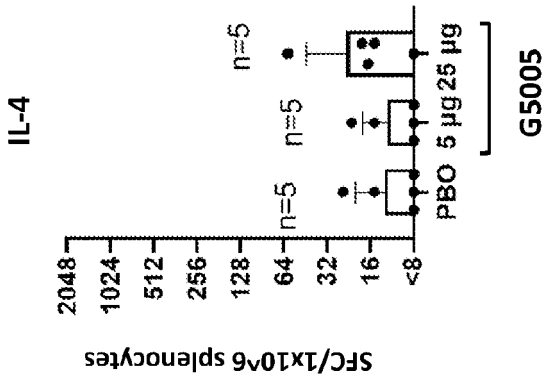

IL-4
SFC/1x10^6 splenocytes
2048
1024
512
256
128
64
32
16
<8
n=5
n=5
n=5
PBO
5 µg
25 µg
G5005

FIG. 5B

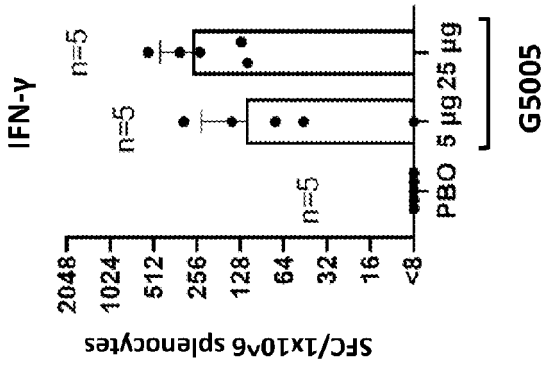

IFN-γ
SFC/1x10^6 splenocytes
2048
1024
512
256
128
64
32
16
<8
n=5
n=5
n=5
PBO
5 µg
25 µg
G5005

FIG. 6A

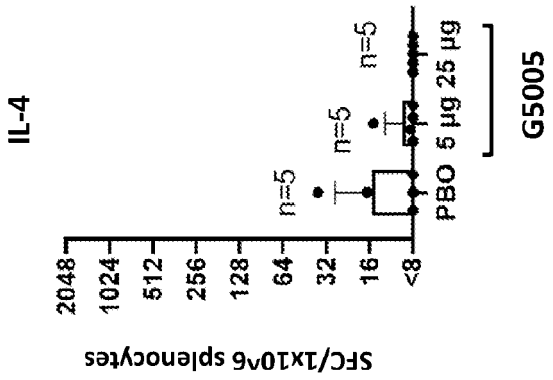

IL-4
SFC/1x10^6 splenocytes
2048
1024
512
256
128
64
32
16
<8
n=5
n=5
n=5
PBO
5 µg
25 µg
G5005

FIG. 6B

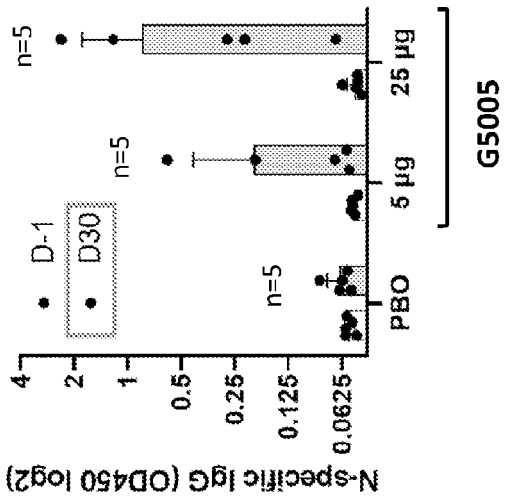

N-specific IgG (OD450 log2)
4
2
1
0.5
0.25
0.125
0.0625
D-1
D30
n=5
n=5
n=5
PBO
5 µg
25 µg
G5005

FIG. 7

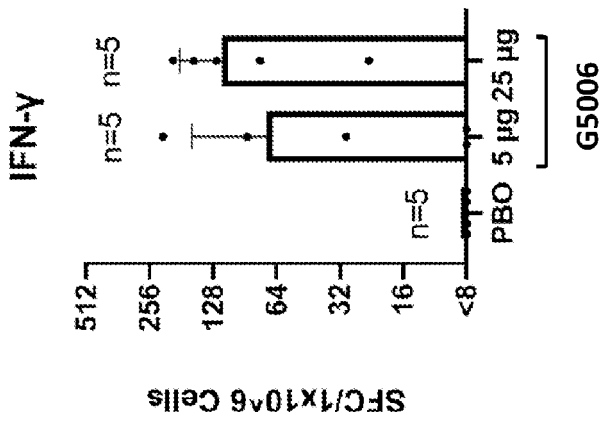

IFN-γ
SFC/1x10^6 Cells
512
256
128
64
32
16
<8
n=5
n=5
n=5
PBO
5 µg
25 µg
G5006

FIG. 8

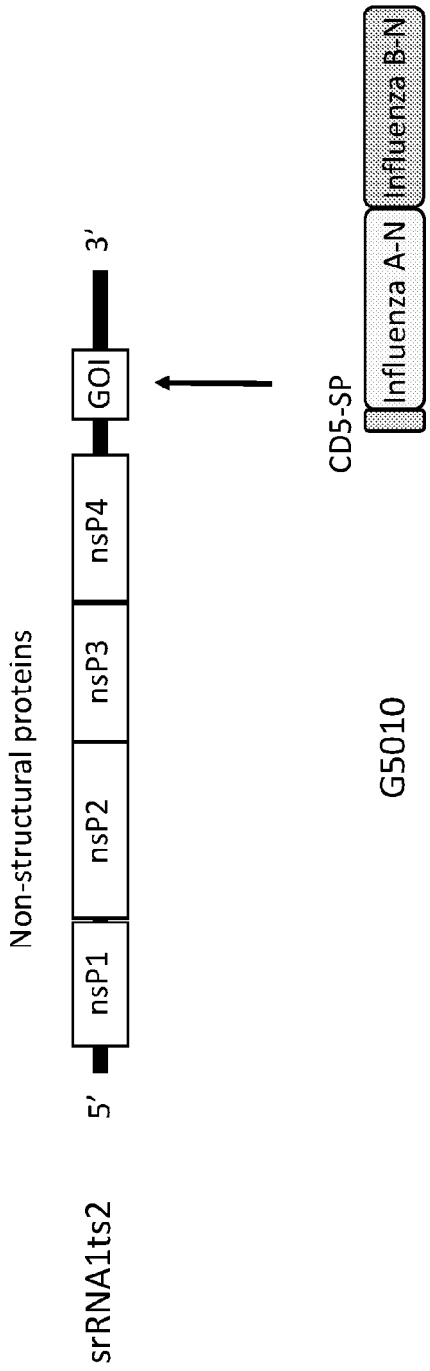

srRNA1ts2
5'
nsP1
nsP2
nsP3
nsP4
GOI
3'
Non-structural proteins
G5010
CD5-SP
Influenza A-N
Influenza B-N

FIG. 9

Query: Nucleoprotein of Influenza A (H5N8 NP; ProteinID AHL21420.1)
Sbjct: Nucleoprotein of Influenza A (AnnArbor H2N2; Prot ID P21433)

```
Query  25   ASQGTKRSYEQMETGGERQNATEIRASVGRMVGGIGRFYIQMCTELKLSDYEGRLIQNSI  84
Sbjct  2    ..............D......N.......K.I..........................L  61

Query  85   TIERMVLSAFDERRNKYLEEHPSAGKDPKKTGGPIYRRRDGKWVRELILYDKEEIRRIWR  144
Sbjct  62   .....................................K.V....M...V.............  121

Query  145  QANNGEDATAGLTHLMIWHSNLNDATYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGA  204
Sbjct  122  .....D........M.........T.................................  181

Query  205  AVKGVGTMVMELIRMIKRGINDRNFWRGENGRRTRIAYERMCNILKGKFQTAAQRAMMDQ  264
Sbjct  182  ................................K..N.........................  241

Query  265  VRESRNPGNAEIEDLIFLARSALILRGSVAHKSCLPACVYGLAVASGYDFEREGYSLVGI  324
Sbjct  242  ........................................P.........K........  301

Query  325  DPFRLLQNSQVFSLIRPNENPAHKSQLVWMACHSAAFEDLRVSSFIRGTRVVPRGQLSTR  384
Sbjct  302  ...K.......Y...................N..............K.I...K....  361

Query  385  GVQIASNENMETMDSSTLELRSRYWAIRTRSGGTTNQQRASAGQISVQPTFSVQRNLPFE  444
Sbjct  362  ..........D..G..................N.......................D  421

Query  445  RATIMAAFTGNTEGRTSDMRTEIIRMMESAKPEDVSFQGRGVFELSDEKATNPIVPSFDM  504
Sbjct  422  KP.........A........A.......G....E.........................  481

Query  505  SNEGSYFFGDNAEEYDN  521
Sbjct  482  .................  498
```

FIG. 10

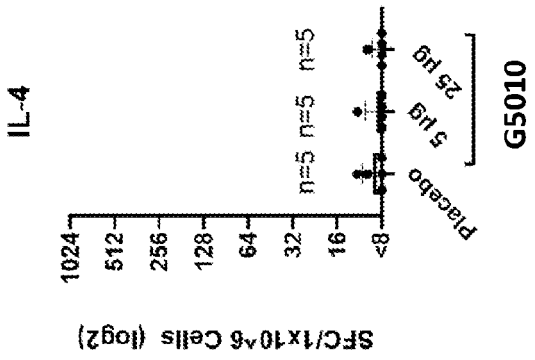


FIG. 11B

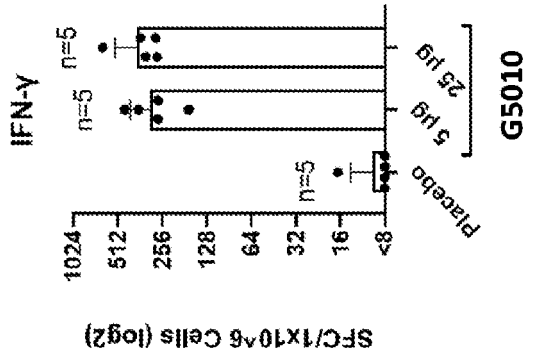


FIG. 11A

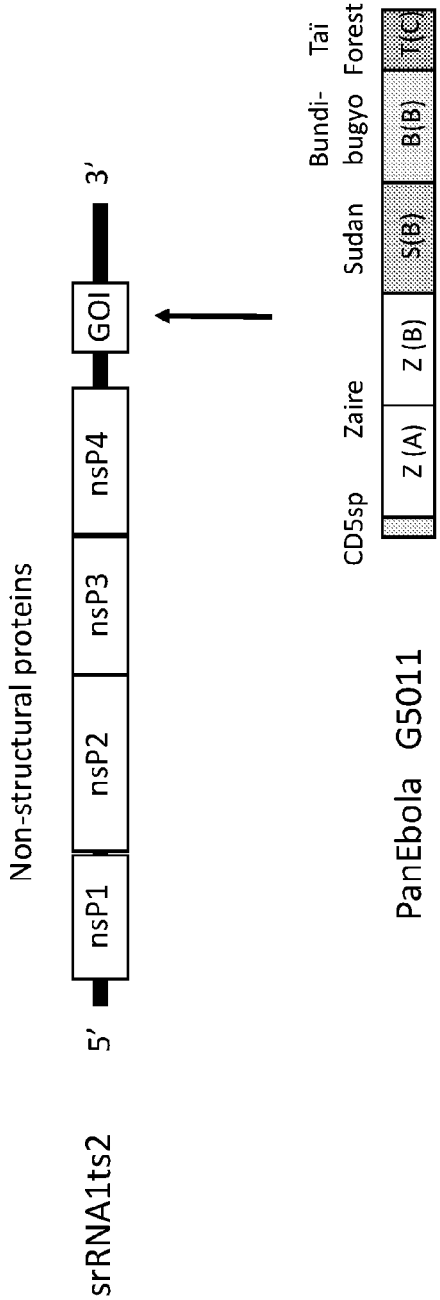

srRNA1ts2
5'
Non-structural proteins
nsP1
nsP2
nsP3
nsP4
GOI
3'
PanEbola G5011
CD5sp
Zaire
Z (A)
Z (B)
Sudan
S(B)
Bundi-bugyo
B(B)
Taï Forest
T(C)

FIG. 12

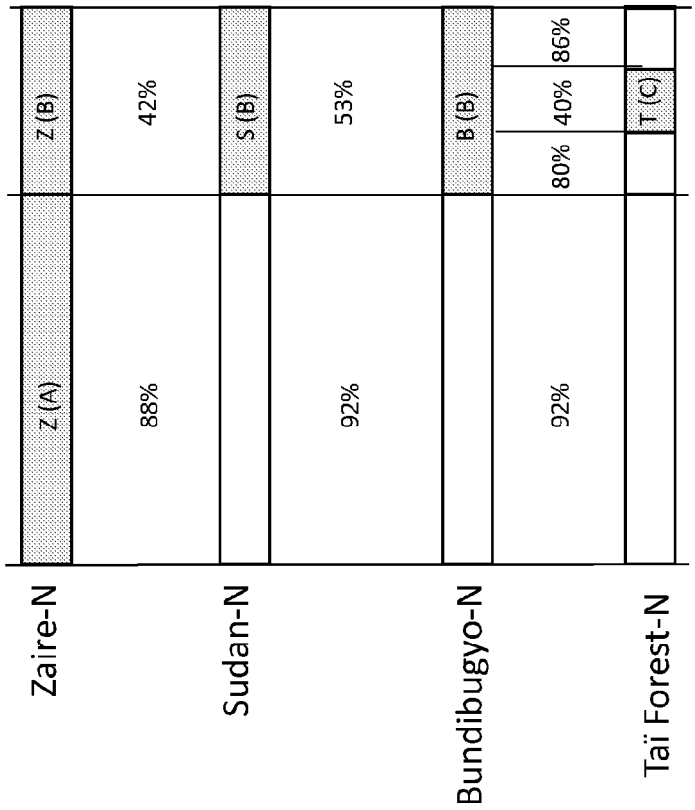

Zaire-N
Z (A)
88%
Z (B)
42%
Sudan-N
92%
S (B)
53%
Bundibugyo-N
92%
B (B)
80%
40%
86%
Taï Forest-N
T (C)

FIG. 13

FIG. 14A

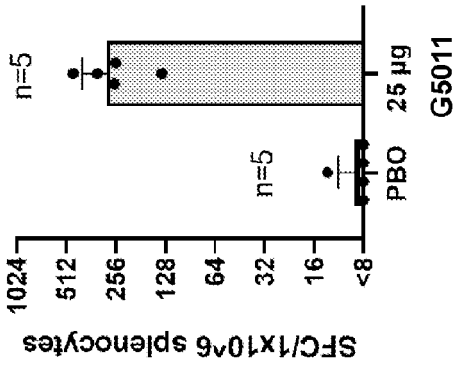
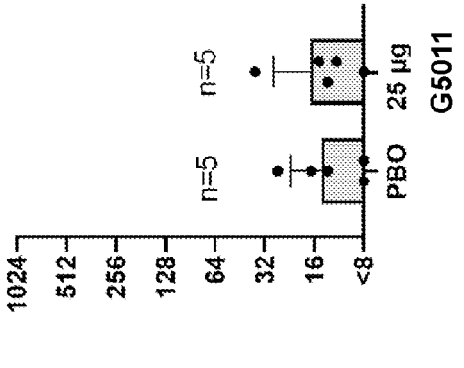


FIG-14B



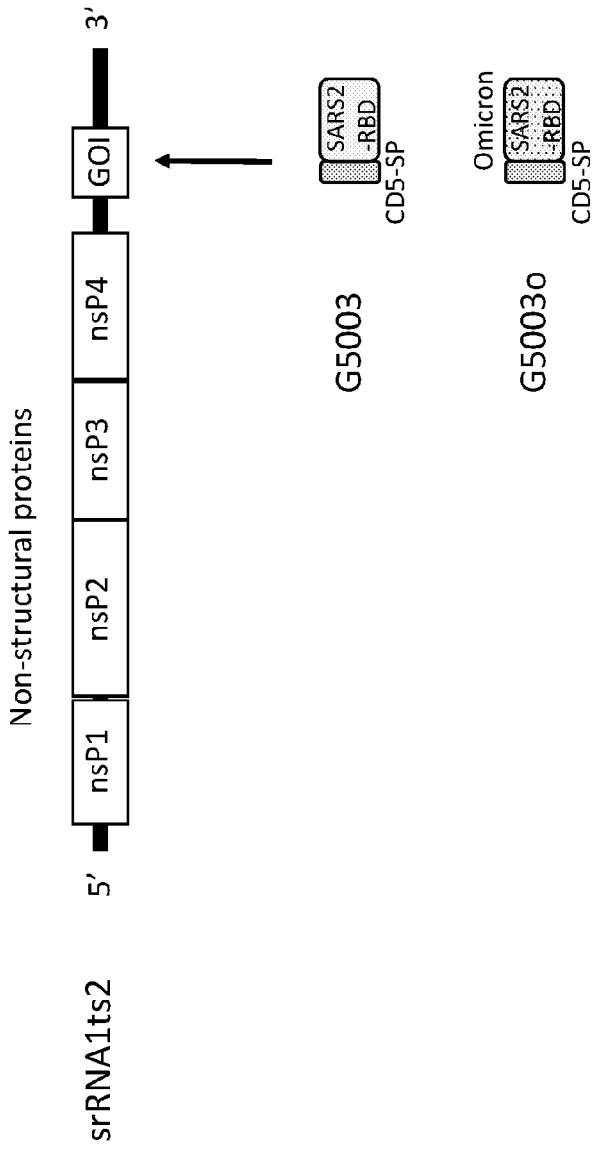

srRNA1ts2
5'
nsP1
nsP2
nsP3
nsP4
GOI
3'
Non-structural proteins
G5003
SARS2
-RBD
CD5-SP
G5003o
Omicron
SARS2
-RBD
CD5-SP

FIG. 15

FIG. 16A

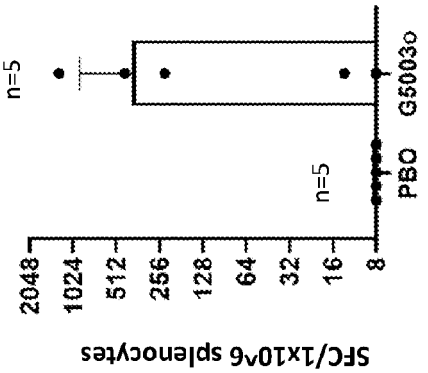


FIG. 16B

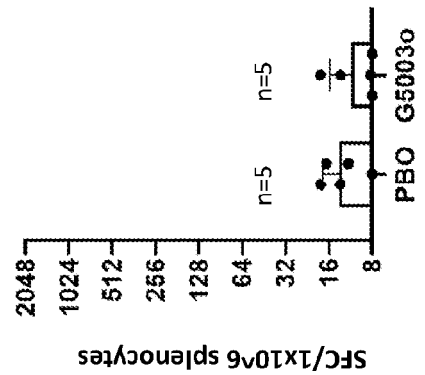


FIG. 17A

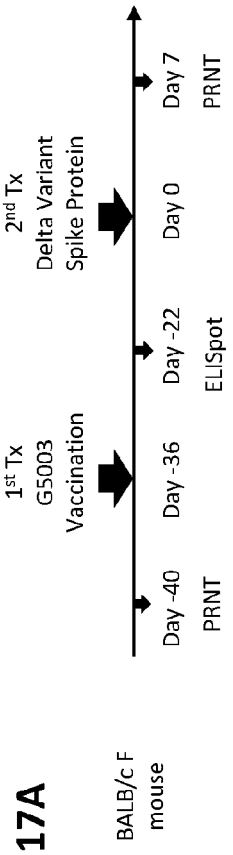


FIG. 17B

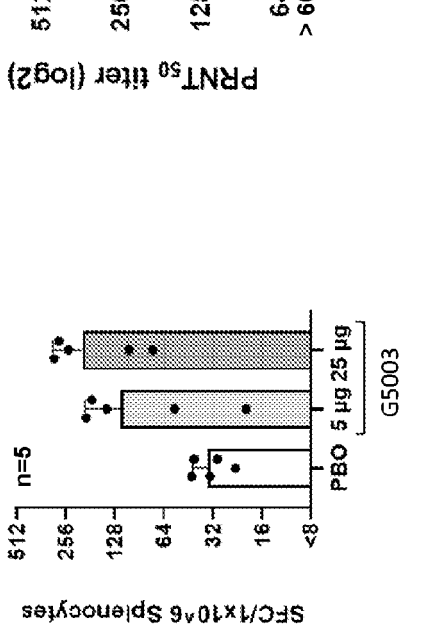


FIG. 17C

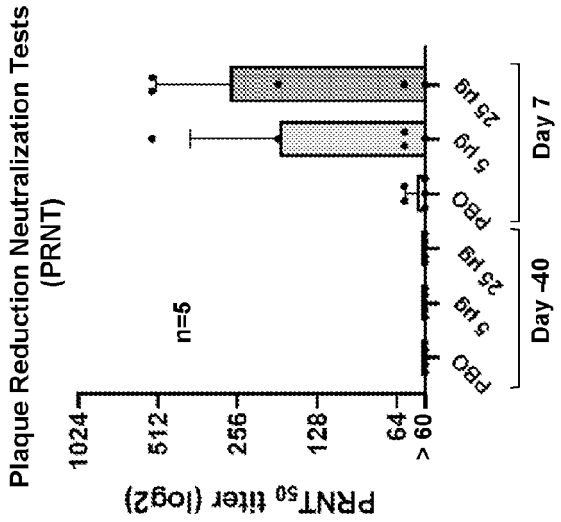


FIG. 18A

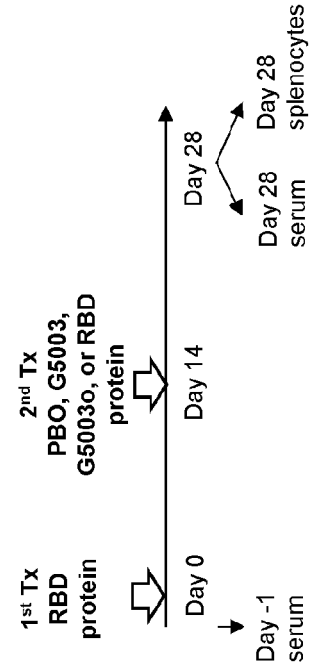


FIG. 18B

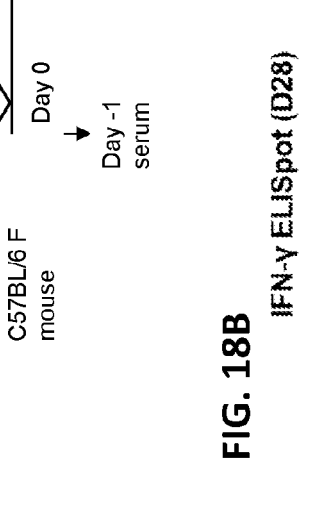
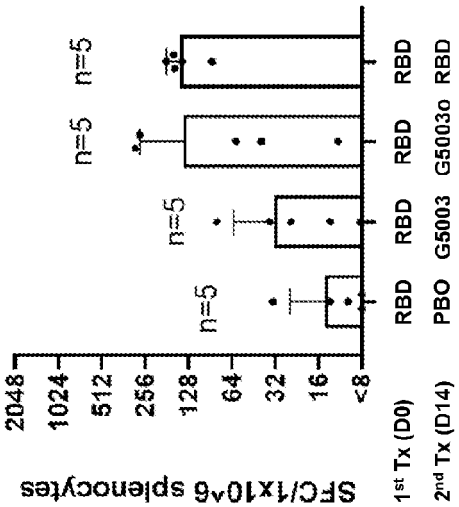


FIG. 18C



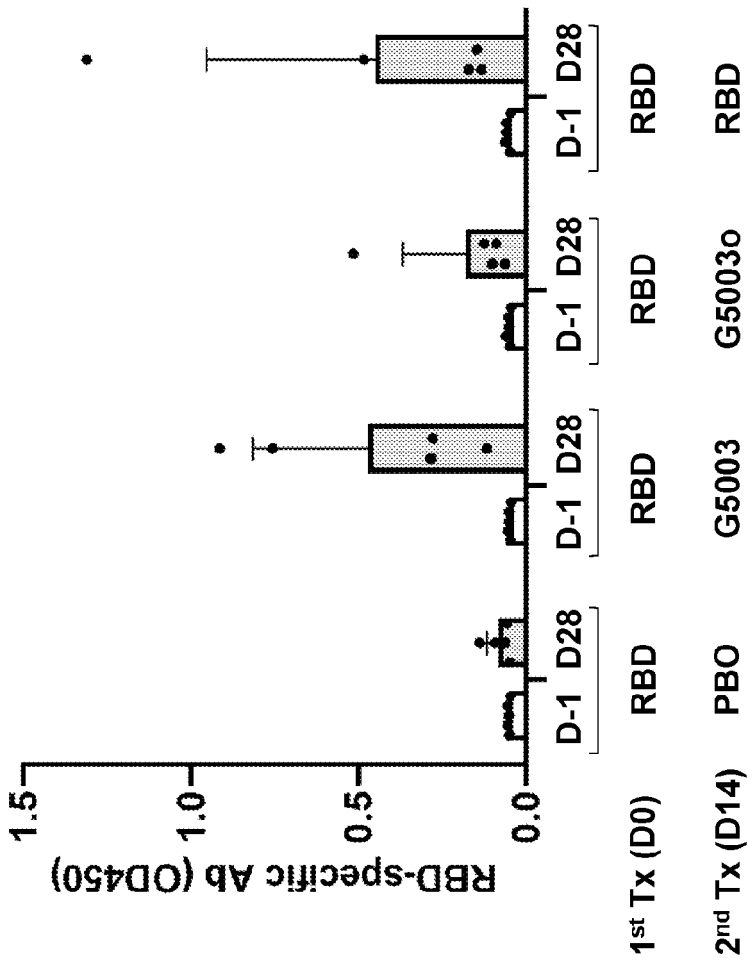

RBD-specific Ab (OD450)
1.5
1.0
0.5
0.0
D-1 D28
D-1 D28
D-1 D28
D-1 D28
1st Tx (D0)
RBD
RBD
RBD
RBD
2nd Tx (D14)
PBO
G5003
G5003o
RBD

FIG. 19

FIG. 20A

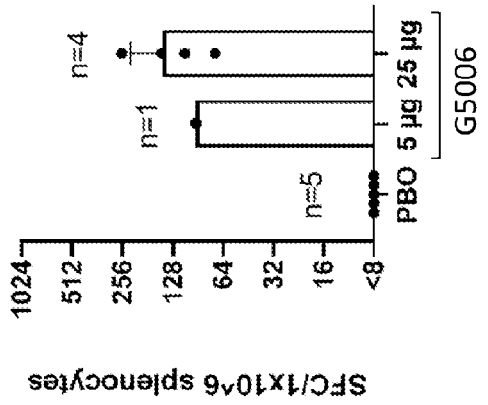


FIG. 20B

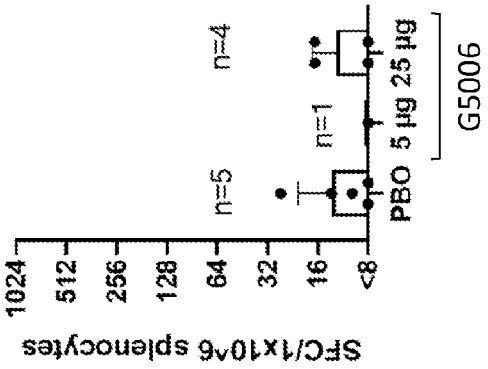


FIG. 20D

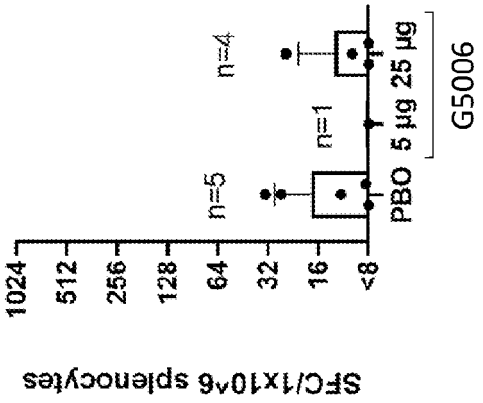
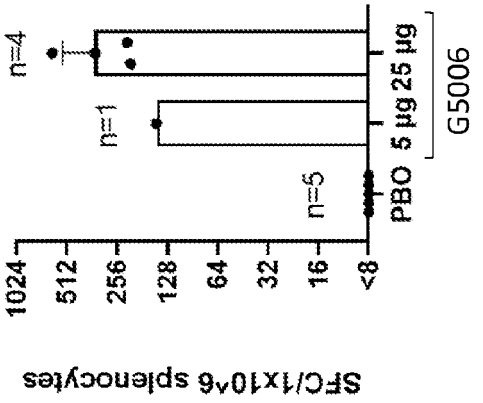


FIG. 20C



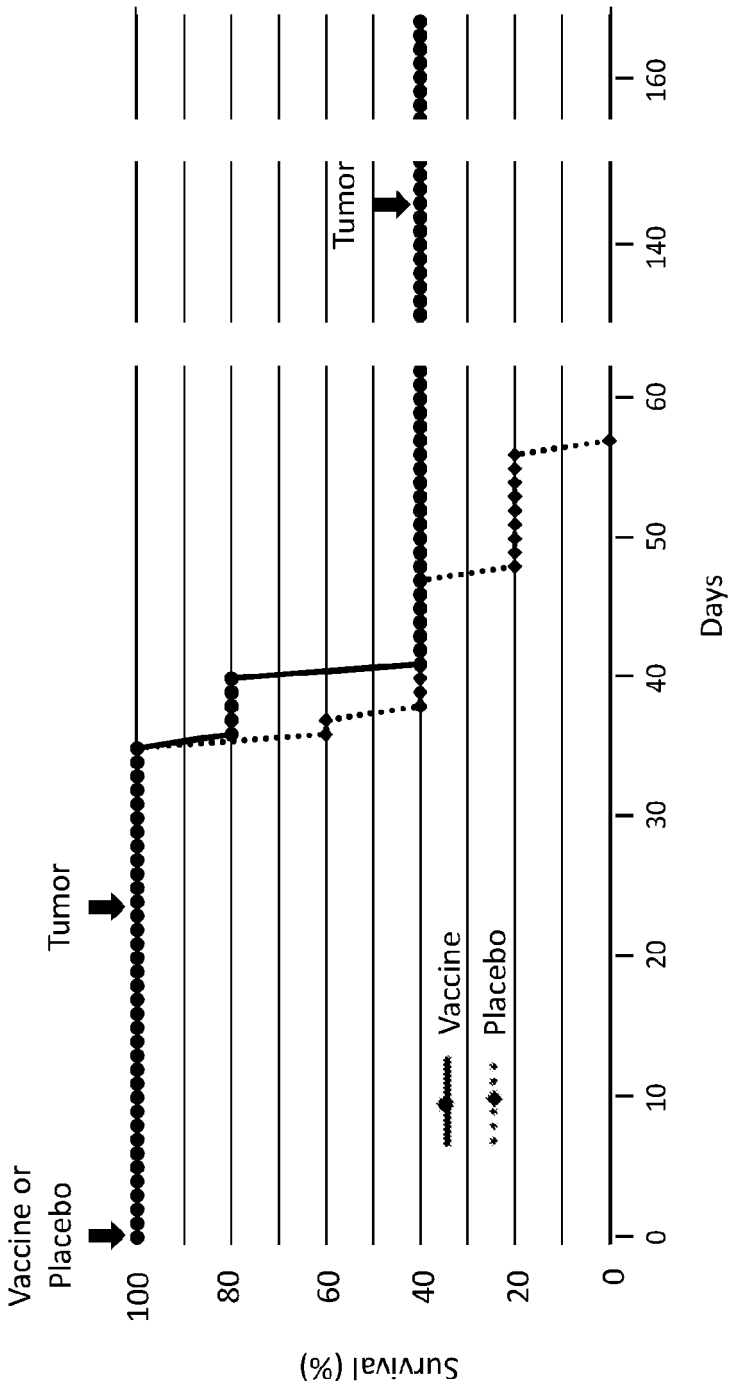

Survival (%)
100
80
60
40
20
0
0
10
20
30
40
50
60
Days
140
160
Vaccine or
Placebo
Tumor
Tumor
Vaccine
Placebo

FIG. 21

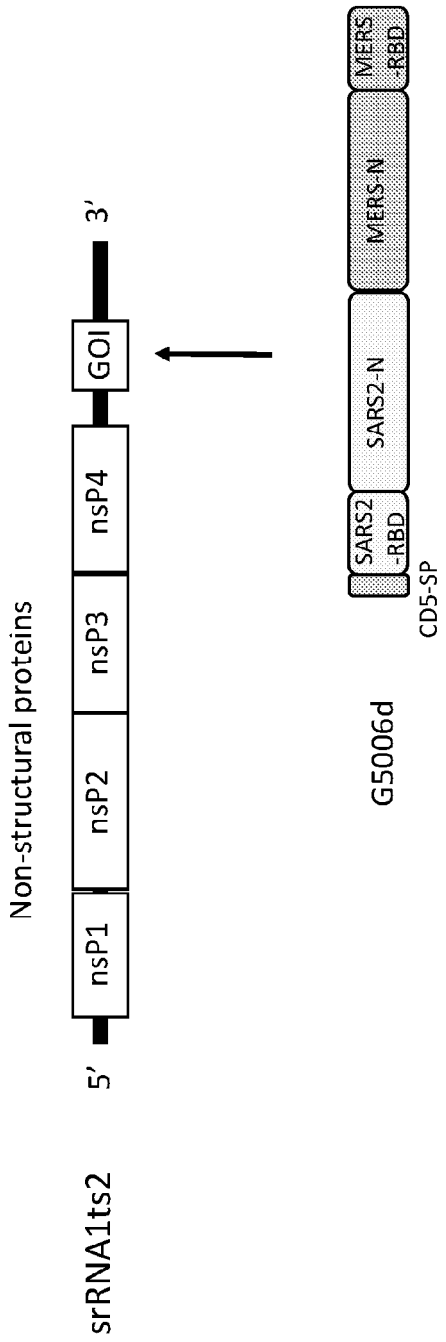

srRNA1ts2
5'
nsP1
nsP2
nsP3
nsP4
GOI
3'
Non-structural proteins
G5006d
CD5-SP
SARS2 -RBD
SARS2-N
MERS-N
MERS -RBD

FIG. 22

FIG. 23A

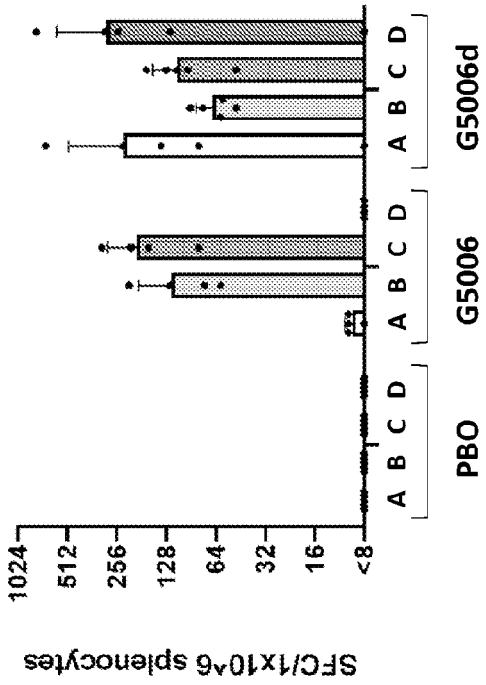
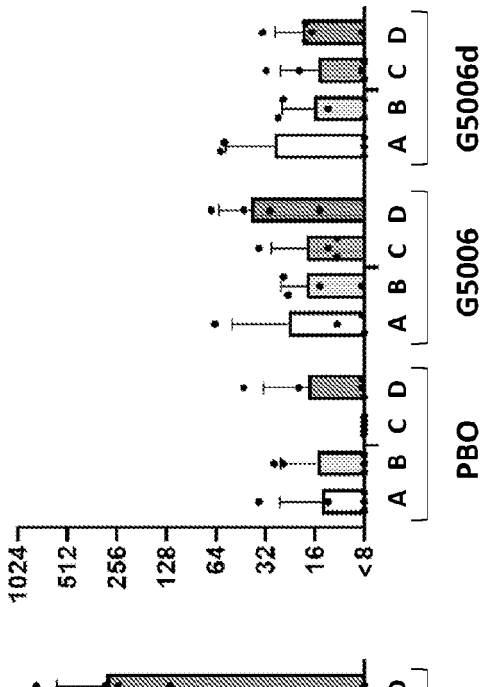


FIG. 23B



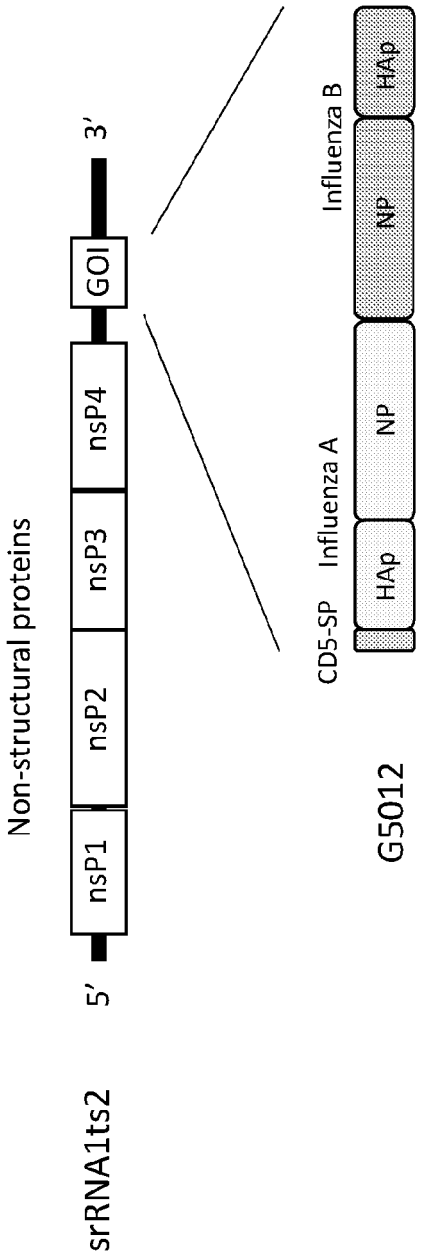

srRNA1ts2
5'
nsP1
nsP2
nsP3
nsP4
GOI
3'
Non-structural proteins
G5012
CD5-SP
Influenza A
HAp
NP
Influenza B
NP
HAp

FIG. 24

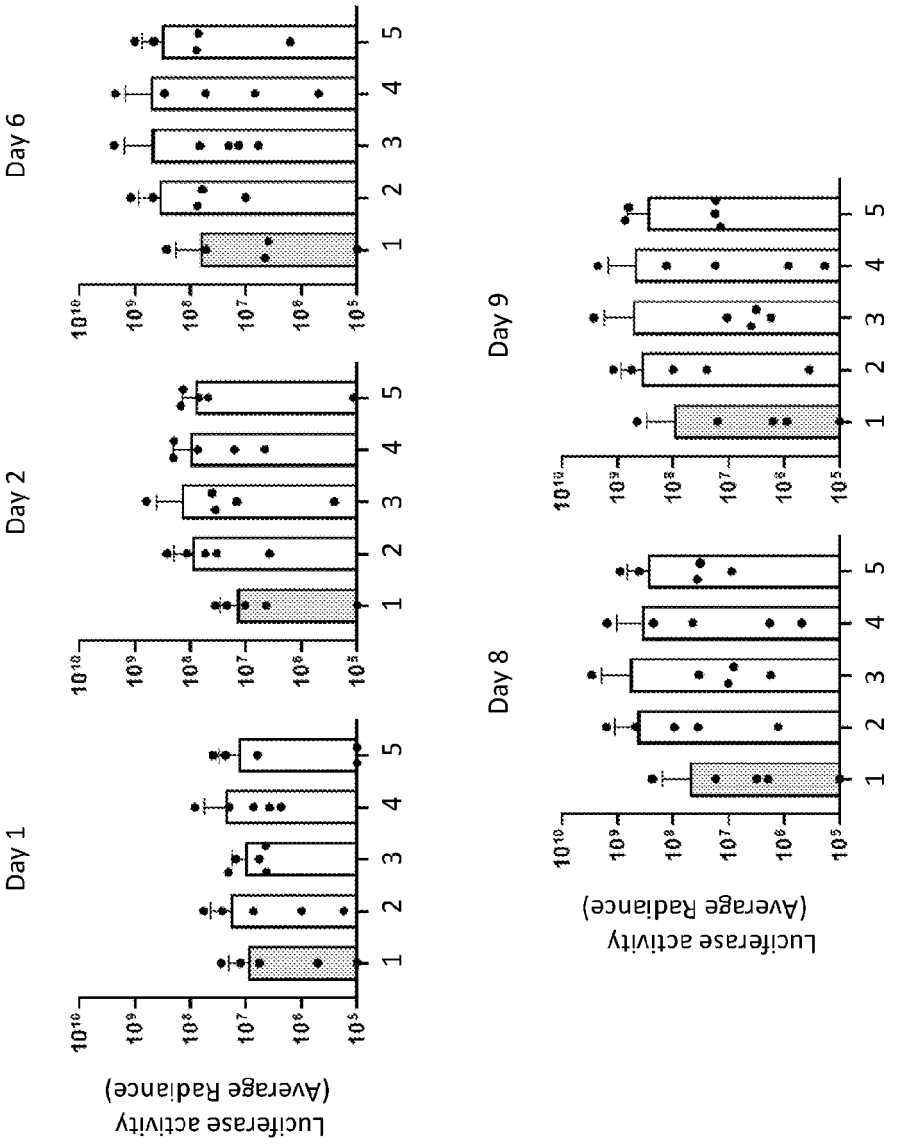

Day 1
Day 2
Day 6
Day 8
Day 9
Luciferase activity
(Average Radiance)
10^{10}
10^{9}
10^{8}
10^{7}
10^{6}
10^{5}
1 2 3 4 5

FIG. 25

SEQUENCE LISTING

<110> Elixirgen Therapeutics, Inc.

<120> TEMPERATURE-CONTROLLABLE, SELF-REPLICATING RNA VACCINES FOR VIRAL DISEASES

<130> 69944-20014.40

<140> Not Yet Assigned
<141> Concurrently Herewith

<150> US 63/275,398
<151> 2021-11-03

<150> US 63/240,278
<151> 2021-09-02

<150> US 63/211,974
<151> 2021-06-17

<160> 29

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1260
<212> DNA
<213> SARS-COV-2

<400> 1
atgtctgata atggacccca aaatcagcga aatgcacccc gcattacgtt tggtggaccc 60
tcagattcaa ctggcagtaa ccagaatgga gaacgcagtg gggcgcgatc aaaacaacgt 120
cggccccaag gtttacccaa taatactgcg tcttggttca ccgctctcac tcaacatggc 180
aaggaagacc ttaaattccc tcgaggacaa ggcgttccaa ttaacaccaa tagcagtcca 240
gatgaccaaa ttggctacta ccgaagagct accagacgaa ttcgtggtgg tgacggtaaa 300
atgaaagatc tcagtccaag atggtatttc tactacctag gaactgggcc agaagctgga 360
cttccctatg gtgctaacaa agacggcatc atatgggttg caactgaggg agccttgaat 420
acaccaaaag atcacattgg cacccgcaat cctgctaaca atgctgcaat cgtgctacaa 480
cttcctcaag gaacaacatt gccaaaaggc ttctacgcag aagggagcag aggcggcagt 540
caagcctctt ctcgttcctc atcacgtagt cgcaacagtt caagaaattc aactccaggc 600
agcagtaggg gaacttctcc tgctagaatg gctggcaatg gcggtgatgc tgctcttgct 660
ttgctgctgc ttgacagatt gaaccagctt gagagcaaaa tgtctggtaa aggccaacaa 720
caacaaggcc aaactgtcac taagaaatct gctgctgagg cttctaagaa gcctcggcaa 780
aaacgtactg ccactaaagc atacaatgta acacaagctt tcggcagacg tggtccagaa 840
caaacccaag gaaattttgg ggaccaggaa ctaatcagac aaggaactga ttacaaacat 900
tggccgcaaa ttgcacaatt tgcccccagc gcttcagcgt tcttcggaat gtcgcgcatt 960
ggcatggaag tcacaccttc gggaacgtgg ttgacctaca caggtgccat caaattggat 1020
gacaaagatc caaatttcaa agatcaagtc attttgctga ataagcatat tgacgcatac 1080

aaaacattcc caccaacaga gcctaaaaag gacaaaaaga agaaggctga tgaaactcaa 1140
gccttaccgc agagacagaa gaaacagcaa actgtgactc ttcttcctgc tgcagatttg 1200
gatgatttct ccaaacaatt gcaacaatcc atgagcagtg ctgactcaac tcaggcctaa 1260

<210> 2
<211> 1329
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 2
atgcccatgg ggtctctgca accgctggcc accttgtacc tgctggggat gctggtcgct 60
tcctgcctcg gatctgataa tggaccccaa aatcagcgaa atgcaccccg cattacgttt 120
ggtggaccct cagattcaac tggcagtaac cagaatggag aacgcagtgg ggcgcgatca 180
aaacaacgtc ggccccaagg tttacccaat aatactgcgt cttggttcac cgctctcact 240
caacatggca aggaagacct taaattccct cgaggacaag gcgttccaat taacaccaat 300
agcagtccag atgaccaaat tggctactac cgaagagcta ccagacgaat tcgtggtggt 360
gacggtaaaa tgaaagatct cagtccaaga tggtatttct actacctagg aactgggcca 420
gaagctggac ttccctatgg tgctaacaaa gacggcatca tatgggttgc aactgaggga 480
gccttgaata caccaaaaga tcacattggc acccgcaatc ctgctaacaa tgctgcaatc 540
gtgctacaac ttcctcaagg aacaacattg ccaaaaggct tctacgcaga agggagcaga 600
ggcggcagtc aagcctcttc tcgttcctca tcacgtagtc gcaacagttc aagaaattca 660
actccaggca gcagtagggg aacttctcct gctagaatgg ctggcaatgg cggtgatgct 720
gctcttgctt tgctgctgct tgacagattg aaccagcttg agagcaaaat gtctggtaaa 780
ggccaacaac aacaaggcca aactgtcact aagaaatctg ctgctgaggc ttctaagaag 840
cctcggcaaa aacgtactgc cactaaagca tacaatgtaa cacaagcttt cggcagacgt 900
ggtccagaac aaacccaagg aaattttggg gaccaggaac taatcagaca aggaactgat 960
tacaaacatt ggccgcaaat tgcacaattt gcccccagcg cttcagcgtt cttcggaatg 1020
tcgcgcattg gcatggaagt cacaccttcg ggaacgtggt tgacctacac aggtgccatc 1080
aaattggatg acaaagatcc aaatttcaaa gatcaagtca ttttgctgaa taagcatatt 1140
gacgcataca aaacattccc accaacagag cctaaaaagg acaaaaagaa gaaggctgat 1200
gaaactcaag ccttaccgca gagacagaag aaacagcaaa ctgtgactct tcttcctgct 1260
gcagatttgg atgatttctc caaacaattg caacaatcca tgagcagtgc tgactcaact 1320
caggcctaa 1329

<210> 3
<211> 2568
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 3
atgcccatgg ggtctctgca accgctggcc accttgtacc tgctggggat gctggtcgct 60
tcctgcctcg gatctgataa tggaccccaa aatcagcgaa atgcaccccg cattacgttt 120
ggtggaccct cagattcaac tggcagtaac cagaatggag aacgcagtgg ggcgcgatca 180

aaacaacgtc ggccccaagg tttacccaat aatactgcgt cttggttcac cgctctcact 240
caacatggca aggaagacct taaattccct cgaggacaag gcgttccaat taacaccaat 300
agcagtccag atgaccaaat tggctactac cgaagagcta ccagacgaat tcgtggtggt 360
gacggtaaaa tgaaagatct cagtccaaga tggtatttct actacctagg aactgggcca 420
gaagctggac ttccctatgg tgctaacaaa gacggcatca tatgggttgc aactgaggga 480
gccttgaata caccaaaaga tcacattggc acccgcaatc ctgctaacaa tgctgcaatc 540
gtgctacaac ttcctcaagg aacaacattg ccaaaaggct tctacgcaga agggagcaga 600
ggcggcagtc aagcctcttc tcgttcctca tcacgtagtc gcaacagttc aagaaattca 660
actccaggca gcagtagggg aacttctcct gctagaatgg ctggcaatgg cggtgatgct 720
gctcttgctt tgctgctgct tgacagattg aaccagcttg agagcaaaat gtctggtaaa 780
ggccaacaac aacaaggcca aactgtcact aagaaatctg ctgctgaggc ttctaagaag 840
cctcggcaaa aacgtactgc cactaaagca tacaatgtaa cacaagcttt cggcagacgt 900
ggtccagaac aaacccaagg aaattttggg gaccaggaac taatcagaca aggaactgat 960
tacaaacatt ggccgcaaat tgcacaattt gcccccagcg cttcagcgtt cttcggaatg 1020
tcgcgcattg gcatggaagt cacaccttcg ggaacgtggt tgacctacac aggtgccatc 1080
aaattggatg acaaagatcc aaatttcaaa gatcaagtca ttttgctgaa taagcatatt 1140
gacgcataca aaacattccc accaacagag cctaaaaagg acaaaaagaa gaaggctgat 1200
gaaactcaag ccttaccgca gagacagaag aaacagcaaa ctgtgactct tcttcctgct 1260
gcagatttgg atgatttctc caaacaattg caacaatcca tgagcagtgc tgactcaact 1320
caggccatgg catcccctgc tgcacctcgt gctgtttcct ttgccgataa caatgatata 1380
acaaatacaa acctatctcg aggtagagga cgtaatccaa aaccacgagc tgcaccaaat 1440
aacactgtct cttggtacac tgggcttacc caacacggga aagtccctct tacctttcca 1500
cctgggcagg gtgtacctct taatgccaat tctacccctg cgcaaaatgc tgggtattgg 1560
cggagacagg acagaaaaat taataccggg aatggaatta agcaactggc tcccaggtgg 1620
tacttctact acactggaac tggacccgaa gcagcactcc cattccgggc tgttaaggat 1680
ggcatcgttt gggtccatga agatggcgcc actgatgctc cttcaacttt tgggacgcgg 1740
aaccctaaca atgattcagc tattgttaca caattcgcgc ccggtactaa gcttcctaaa 1800
aacttccaca ttgaggggac tggaggcaat agtcaatcat cttcaagagc ctctagctta 1860
agcagaaact cttccagatc tagttcacaa ggttcaagat caggaaactc tacccgcggc 1920
acttctccag gtccatctgg aatcggagca gtaggaggtg atctacttta ccttgatctt 1980
ctgaacagac tacaagccct tgagtctggc aaagtaaagc aatcgcagcc aaaagtaatc 2040
actaagaaag atgctgctgc tgctaaaaat aagatgcgcc acaagcgcac ttccaccaaa 2100
agtttcaaca tggtgcaagc ttttggtctt cgcggaccag gagacctcca gggaaacttt 2160
ggtgatcttc aattgaataa actcggcact gaggacccac gttggcccca aattgctgag 2220
cttgctccta cagccagtgc ttttatgggt atgtcgcaat ttaaacttac ccatcagaac 2280
aatgatgatc atggcaaccc tgtgtacttc cttcggtaca gtggagccat taaacttgac 2340
ccaaagaatc ccaactacaa taagtggttg gagcttcttg agcaaaatat tgatgcctac 2400
aaaaccttcc ctaagaagga aaagaaacaa aaggcaccaa aagaagaatc aacagaccaa 2460
atgtctgaac ctccaaagga gcagcgtgtg caaggtagca tcactcagcg cactcgcacc 2520
cgtccaagtg ttcagcctgg tccaatgatt gatgttaaca ctgattag 2568

<210> 4
<211> 2568
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 4

atgcctatgg gctctctgca gcccctggcc accctgtacc tgctgggcat gctggtggcc 60
agctgcctgg gaagcgacaa cggcccccag aaccagagaa acgcccctag aatcacattt 120
ggcggcccta gtgatagcac cggatctaat caaaacggcg agagaagcgg cgctcggtct 180
aaacagagac ggccacaggg actgcctaac aacaccgcca gctggttcac cgccctgacc 240
cagcacggca aggaggacct taagttccct cggggacagg gcgtgccaat caacaccaac 300
tctagtcccg acgaccagat cggctattat agaagagcca caagacgcat cagaggtggc 360
gacggcaaga tgaaggacct gagccctcgc tggtactttt actacctggg gaccggccct 420
gaagccggcc tgccttacgg cgccaacaag gacggaatca tctgggtcgc caccgagggc 480
gccctgaata cccctaagga ccacatcggc accagaaacc ctgctaataa tgccgctatc 540
gtgctgcagc tgcctcaggg caccaccctg cctaagggct tctacgccga gggctcccgg 600
ggaggttccc aggctagcag cagatcttcc agccggagca gaaacagctc caggaacagc 660
acacctggca gcagcagagg tacgagccct gcccggatgg ccggaaacgg cggcgatgcc 720
gccctggccc tgctgctgct ggacagactg aaccagctcg agagcaagat gtctggcaag 780
ggccagcagc agcagggcca gacagtgacc aagaaatccg ccgctgaggc cagcaaaaaa 840
cccagacaga aaagaaccgc tacaaaggcc tacaacgtta cccaggcctt tggcagacgg 900
ggcccagagc agacccaggg aaacttcggc gaccaggagc tgatccggca gggcacggac 960
tacaagcact ggcctcaaat cgcccagttt gccccttccg ccagcgcttt cttcggaatg 1020
agcagaatcg gcatggaagt gacacctagt ggcacctggc tgacctacac cggtgccatt 1080
aagctggatg acaaggaccc caacttcaag gatcaggtga tcctgctgaa caagcacatt 1140
gatgcttaca agaccttccc acctaccgag ccaaaaaaag ataagaagaa aaaagccgat 1200
gagacacaag ccctgcccca gaggcagaag aagcaacaaa ccgtcaccct gctgcctgct 1260
gccgacctgg acgacttcag caaacagctg cagcagagca tgagctctgc tgatagcacc 1320
caggccatgg cctctccagc cgctcccaga gctgtgtcct tcgccgataa taacgacatc 1380
acaaacacca acctgagccg gggcagaggc agaaacccta aacctagagc cgcccccaac 1440
aacaccgtga gctggtatac aggcctcacc cagcatggca aggtgcctct gacattcccc 1500
cctgggcagg gcgtgcccct gaacgccaac agcacccctg cccagaatgc cggctactgg 1560
cggaggcagg acagaaagat caacactggt aacggcatca agcagctggc cccacggtgg 1620
tatttctact acaccggcac cggcccggaa gccgccctcc ccttcagagc cgtgaaggac 1680
ggcatcgtgt gggtgcacga ggacggcgcc acagatgccc cgtctacatt tggcactcgg 1740
aatcccaata acgacagcgc catcgtgacc cagttcgccc ctggcaccaa gctgcctaag 1800
aactttcaca tcgagggcac aggaggcaac agccagagca gcagccgggc ttcgagcctg 1860
tctcggaata gctcccggtc cagctctcag ggcagccgca gtggaaattc cacccggggc 1920
acatctcctg gccccagcgg catcggcgct gtgggcggag acctgctcta cctggacctg 1980
ctgaacagac tgcaggcact tgaaagcggc aaagttaagc aatctcaacc taaggtgatc 2040
accaaaaagg acgccgccgc cgctaagaac aagatgagac acaagagaac aagcacaaag 2100
agcttcaaca tggtgcaagc cttcggcctg cggggacctg gcgacctgca gggcaacttc 2160
ggcgacctgc agctgaacaa gctgggcaca gaggatcctc gatggcccca gatcgccgaa 2220
ctagctccaa ccgccagcgc cttcatgggc atgagccagt tcaagctgac acaccagaac 2280
aatgacgatc acggaaatcc tgtgtacttc ctgagataca gcggcgccat caagctggat 2340
cctaagaacc ccaactacaa caagtggctg gaactgctgg aacagaacat cgacgcctac 2400
aagaccttcc ccaagaagga aaagaagcag aaggccccta aagaggaaag cacagatcag 2460
atgagcgagc ctcccaagga acagagagtg cagggatcta tcacccagag aacaagaaca 2520
agacccagcg tgcagcctgg ccctatgatt gacgtgaaca ccgactag 2568

<210> 5
<211> 419
<212> PRT
<213> SARS-COV-2

<400> 5

```
Met Ser Asp Asn Gly Pro Gln Asn Gln Arg Asn Ala Pro Arg Ile Thr
 1               5                  10                  15
Phe Gly Gly Pro Ser Asp Ser Thr Gly Ser Asn Gln Asn Gly Glu Arg
            20                  25                  30
Ser Gly Ala Arg Ser Lys Gln Arg Arg Pro Gln Gly Leu Pro Asn Asn
        35                  40                  45
Thr Ala Ser Trp Phe Thr Ala Leu Thr Gln His Gly Lys Glu Asp Leu
    50                  55                  60
Lys Phe Pro Arg Gly Gln Gly Val Pro Ile Asn Thr Asn Ser Ser Pro
65                  70                  75                  80
Asp Asp Gln Ile Gly Tyr Tyr Arg Arg Ala Thr Arg Arg Ile Arg Gly
                85                  90                  95
Gly Asp Gly Lys Met Lys Asp Leu Ser Pro Arg Trp Tyr Phe Tyr Tyr
            100                 105                 110
Leu Gly Thr Gly Pro Glu Ala Gly Leu Pro Tyr Gly Ala Asn Lys Asp
        115                 120                 125
Gly Ile Ile Trp Val Ala Thr Glu Gly Ala Leu Asn Thr Pro Lys Asp
    130                 135                 140
His Ile Gly Thr Arg Asn Pro Ala Asn Asn Ala Ala Ile Val Leu Gln
145                 150                 155                 160
Leu Pro Gln Gly Thr Thr Leu Pro Lys Gly Phe Tyr Ala Glu Gly Ser
                165                 170                 175
Arg Gly Gly Ser Gln Ala Ser Ser Arg Ser Ser Ser Arg Ser Arg Asn
            180                 185                 190
Ser Ser Arg Asn Ser Thr Pro Gly Ser Ser Arg Gly Thr Ser Pro Ala
        195                 200                 205
Arg Met Ala Gly Asn Gly Gly Asp Ala Ala Leu Ala Leu Leu Leu Leu
    210                 215                 220
Asp Arg Leu Asn Gln Leu Glu Ser Lys Met Ser Gly Lys Gly Gln Gln
225                 230                 235                 240
Gln Gln Gly Gln Thr Val Thr Lys Lys Ser Ala Ala Glu Ala Ser Lys
                245                 250                 255
Lys Pro Arg Gln Lys Arg Thr Ala Thr Lys Ala Tyr Asn Val Thr Gln
            260                 265                 270
Ala Phe Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly Asp
        275                 280                 285
Gln Glu Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln Ile
    290                 295                 300
Ala Gln Phe Ala Pro Ser Ala Ser Ala Phe Phe Gly Met Ser Arg Ile
305                 310                 315                 320
Gly Met Glu Val Thr Pro Ser Gly Thr Trp Leu Thr Tyr Thr Gly Ala
                325                 330                 335
Ile Lys Leu Asp Asp Lys Asp Pro Asn Phe Lys Asp Gln Val Ile Leu
            340                 345                 350
Leu Asn Lys His Ile Asp Ala Tyr Lys Thr Phe Pro Pro Thr Glu Pro
        355                 360                 365
Lys Lys Asp Lys Lys Lys Lys Ala Asp Glu Thr Gln Ala Leu Pro Gln
    370                 375                 380
Arg Gln Lys Lys Gln Gln Thr Val Thr Leu Leu Pro Ala Ala Asp Leu
385                 390                 395                 400
```

```
Asp Asp Phe Ser Lys Gln Leu Gln Gln Ser Met Ser Ser Ala Asp Ser
                405                 410                 415
Thr Gln Ala
```

<210> 6
<211> 442
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 6

```
Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1               5                  10                  15
Met Leu Val Ala Ser Cys Leu Gly Ser Asp Asn Gly Pro Gln Asn Gln
            20                  25                  30
Arg Asn Ala Pro Arg Ile Thr Phe Gly Gly Pro Ser Asp Ser Thr Gly
        35                  40                  45
Ser Asn Gln Asn Gly Glu Arg Ser Gly Ala Arg Ser Lys Gln Arg Arg
    50                  55                  60
Pro Gln Gly Leu Pro Asn Asn Thr Ala Ser Trp Phe Thr Ala Leu Thr
65                  70                  75                  80
Gln His Gly Lys Glu Asp Leu Lys Phe Pro Arg Gly Gln Gly Val Pro
                85                  90                  95
Ile Asn Thr Asn Ser Ser Pro Asp Asp Gln Ile Gly Tyr Tyr Arg Arg
            100                 105                 110
Ala Thr Arg Arg Ile Arg Gly Gly Asp Gly Lys Met Lys Asp Leu Ser
        115                 120                 125
Pro Arg Trp Tyr Phe Tyr Tyr Leu Gly Thr Gly Pro Glu Ala Gly Leu
    130                 135                 140
Pro Tyr Gly Ala Asn Lys Asp Gly Ile Ile Trp Val Ala Thr Glu Gly
145                 150                 155                 160
Ala Leu Asn Thr Pro Lys Asp His Ile Gly Thr Arg Asn Pro Ala Asn
                165                 170                 175
Asn Ala Ala Ile Val Leu Gln Leu Pro Gln Gly Thr Thr Leu Pro Lys
            180                 185                 190
Gly Phe Tyr Ala Glu Gly Ser Arg Gly Gly Ser Gln Ala Ser Ser Arg
        195                 200                 205
Ser Ser Ser Arg Ser Arg Asn Ser Ser Arg Asn Ser Thr Pro Gly Ser
    210                 215                 220
Ser Arg Gly Thr Ser Pro Ala Arg Met Ala Gly Asn Gly Gly Asp Ala
225                 230                 235                 240
Ala Leu Ala Leu Leu Leu Leu Asp Arg Leu Asn Gln Leu Glu Ser Lys
                245                 250                 255
Met Ser Gly Lys Gly Gln Gln Gln Gln Gly Gln Thr Val Thr Lys Lys
            260                 265                 270
Ser Ala Ala Glu Ala Ser Lys Lys Pro Arg Gln Lys Arg Thr Ala Thr
```

```
        275                 280                 285
Lys Ala Tyr Asn Val Thr Gln Ala Phe Gly Arg Arg Gly Pro Glu Gln
    290                 295                 300
Thr Gln Gly Asn Phe Gly Asp Gln Glu Leu Ile Arg Gln Gly Thr Asp
305                 310                 315                 320
Tyr Lys His Trp Pro Gln Ile Ala Gln Phe Ala Pro Ser Ala Ser Ala
                325                 330                 335
Phe Phe Gly Met Ser Arg Ile Gly Met Glu Val Thr Pro Ser Gly Thr
            340                 345                 350
Trp Leu Thr Tyr Thr Gly Ala Ile Lys Leu Asp Asp Lys Asp Pro Asn
        355                 360                 365
Phe Lys Asp Gln Val Ile Leu Leu Asn Lys His Ile Asp Ala Tyr Lys
    370                 375                 380
Thr Phe Pro Pro Thr Glu Pro Lys Lys Asp Lys Lys Lys Lys Ala Asp
385                 390                 395                 400
Glu Thr Gln Ala Leu Pro Gln Arg Gln Lys Lys Gln Gln Thr Val Thr
                405                 410                 415
Leu Leu Pro Ala Ala Asp Leu Asp Asp Phe Ser Lys Gln Leu Gln Gln
            420                 425                 430
Ser Met Ser Ser Ala Asp Ser Thr Gln Ala
        435                 440


<210> 7
<211> 855
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 7
Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1               5                  10                  15
Met Leu Val Ala Ser Cys Leu Gly Ser Asp Asn Gly Pro Gln Asn Gln
            20                  25                  30
Arg Asn Ala Pro Arg Ile Thr Phe Gly Gly Pro Ser Asp Ser Thr Gly
        35                  40                  45
Ser Asn Gln Asn Gly Glu Arg Ser Gly Ala Arg Ser Lys Gln Arg Arg
    50                  55                  60
Pro Gln Gly Leu Pro Asn Asn Thr Ala Ser Trp Phe Thr Ala Leu Thr
65                  70                  75                  80
Gln His Gly Lys Glu Asp Leu Lys Phe Pro Arg Gly Gln Gly Val Pro
                85                  90                  95
Ile Asn Thr Asn Ser Ser Pro Asp Asp Gln Ile Gly Tyr Tyr Arg Arg
            100                 105                 110
Ala Thr Arg Arg Ile Arg Gly Gly Asp Gly Lys Met Lys Asp Leu Ser
        115                 120                 125
Pro Arg Trp Tyr Phe Tyr Tyr Leu Gly Thr Gly Pro Glu Ala Gly Leu
    130                 135                 140
```

```
Pro Tyr Gly Ala Asn Lys Asp Gly Ile Ile Trp Val Ala Thr Glu Gly
145                 150                 155                 160
Ala Leu Asn Thr Pro Lys Asp His Ile Gly Thr Arg Asn Pro Ala Asn
                165                 170                 175
Asn Ala Ala Ile Val Leu Gln Leu Pro Gln Gly Thr Thr Leu Pro Lys
            180                 185                 190
Gly Phe Tyr Ala Glu Gly Ser Arg Gly Gly Ser Gln Ala Ser Ser Arg
        195                 200                 205
Ser Ser Ser Arg Ser Arg Asn Ser Ser Arg Asn Ser Thr Pro Gly Ser
    210                 215                 220
Ser Arg Gly Thr Ser Pro Ala Arg Met Ala Gly Asn Gly Gly Asp Ala
225                 230                 235                 240
Ala Leu Ala Leu Leu Leu Leu Asp Arg Leu Asn Gln Leu Glu Ser Lys
                245                 250                 255
Met Ser Gly Lys Gly Gln Gln Gln Gln Gly Gln Thr Val Thr Lys Lys
            260                 265                 270
Ser Ala Ala Glu Ala Ser Lys Lys Pro Arg Gln Lys Arg Thr Ala Thr
        275                 280                 285
Lys Ala Tyr Asn Val Thr Gln Ala Phe Gly Arg Arg Gly Pro Glu Gln
    290                 295                 300
Thr Gln Gly Asn Phe Gly Asp Gln Glu Leu Ile Arg Gln Gly Thr Asp
305                 310                 315                 320
Tyr Lys His Trp Pro Gln Ile Ala Gln Phe Ala Pro Ser Ala Ser Ala
                325                 330                 335
Phe Phe Gly Met Ser Arg Ile Gly Met Glu Val Thr Pro Ser Gly Thr
            340                 345                 350
Trp Leu Thr Tyr Thr Gly Ala Ile Lys Leu Asp Asp Lys Asp Pro Asn
        355                 360                 365
Phe Lys Asp Gln Val Ile Leu Leu Asn Lys His Ile Asp Ala Tyr Lys
    370                 375                 380
Thr Phe Pro Pro Thr Glu Pro Lys Lys Asp Lys Lys Lys Lys Ala Asp
385                 390                 395                 400
Glu Thr Gln Ala Leu Pro Gln Arg Gln Lys Lys Gln Gln Thr Val Thr
                405                 410                 415
Leu Leu Pro Ala Ala Asp Leu Asp Asp Phe Ser Lys Gln Leu Gln Gln
            420                 425                 430
Ser Met Ser Ser Ala Asp Ser Thr Gln Ala Met Ala Ser Pro Ala Ala
        435                 440                 445
Pro Arg Ala Val Ser Phe Ala Asp Asn Asn Asp Ile Thr Asn Thr Asn
    450                 455                 460
Leu Ser Arg Gly Arg Gly Arg Asn Pro Lys Pro Arg Ala Ala Pro Asn
465                 470                 475                 480
Asn Thr Val Ser Trp Tyr Thr Gly Leu Thr Gln His Gly Lys Val Pro
                485                 490                 495
Leu Thr Phe Pro Pro Gly Gln Gly Val Pro Leu Asn Ala Asn Ser Thr
            500                 505                 510
Pro Ala Gln Asn Ala Gly Tyr Trp Arg Arg Gln Asp Arg Lys Ile Asn
        515                 520                 525
Thr Gly Asn Gly Ile Lys Gln Leu Ala Pro Arg Trp Tyr Phe Tyr Tyr
    530                 535                 540
```

```
Thr Gly Thr Gly Pro Glu Ala Ala Leu Pro Phe Arg Ala Val Lys Asp
545                 550                 555                 560
Gly Ile Val Trp Val His Glu Asp Gly Ala Thr Asp Ala Pro Ser Thr
                565                 570                 575
Phe Gly Thr Arg Asn Pro Asn Asn Asp Ser Ala Ile Val Thr Gln Phe
            580                 585                 590
Ala Pro Gly Thr Lys Leu Pro Lys Asn Phe His Ile Glu Gly Thr Gly
        595                 600                 605
Gly Asn Ser Gln Ser Ser Ser Arg Ala Ser Ser Leu Ser Arg Asn Ser
    610                 615                 620
Ser Arg Ser Ser Ser Gln Gly Ser Arg Ser Gly Asn Ser Thr Arg Gly
625                 630                 635                 640
Thr Ser Pro Gly Pro Ser Gly Ile Gly Ala Val Gly Gly Asp Leu Leu
                645                 650                 655
Tyr Leu Asp Leu Leu Asn Arg Leu Gln Ala Leu Glu Ser Gly Lys Val
            660                 665                 670
Lys Gln Ser Gln Pro Lys Val Ile Thr Lys Lys Asp Ala Ala Ala Ala
        675                 680                 685
Lys Asn Lys Met Arg His Lys Arg Thr Ser Thr Lys Ser Phe Asn Met
    690                 695                 700
Val Gln Ala Phe Gly Leu Arg Gly Pro Gly Asp Leu Gln Gly Asn Phe
705                 710                 715                 720
Gly Asp Leu Gln Leu Asn Lys Leu Gly Thr Glu Asp Pro Arg Trp Pro
                725                 730                 735
Gln Ile Ala Glu Leu Ala Pro Thr Ala Ser Ala Phe Met Gly Met Ser
            740                 745                 750
Gln Phe Lys Leu Thr His Gln Asn Asn Asp Asp His Gly Asn Pro Val
        755                 760                 765
Tyr Phe Leu Arg Tyr Ser Gly Ala Ile Lys Leu Asp Pro Lys Asn Pro
    770                 775                 780
Asn Tyr Asn Lys Trp Leu Glu Leu Leu Glu Gln Asn Ile Asp Ala Tyr
785                 790                 795                 800
Lys Thr Phe Pro Lys Lys Glu Lys Lys Gln Lys Ala Pro Lys Glu Glu
                805                 810                 815
Ser Thr Asp Gln Met Ser Glu Pro Pro Lys Glu Gln Arg Val Gln Gly
            820                 825                 830
Ser Ile Thr Gln Arg Thr Arg Thr Arg Pro Ser Val Gln Pro Gly Pro
        835                 840                 845
Met Ile Asp Val Asn Thr Asp
    850                 855


<210> 8
<211> 24
<212> PRT
<213> Homo sapiens

<400> 8
Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1               5                  10                  15
```

```
Met Leu Val Ala Ser Cys Leu Gly
            20


<210> 9
<211> 413
<212> PRT
<213> MERS

<400> 9
Met Ala Ser Pro Ala Ala Pro Arg Ala Val Ser Phe Ala Asp Asn Asn
 1               5                  10                  15
Asp Ile Thr Asn Thr Asn Leu Ser Arg Gly Arg Gly Arg Asn Pro Lys
            20                  25                  30
Pro Arg Ala Ala Pro Asn Asn Thr Val Ser Trp Tyr Thr Gly Leu Thr
        35                  40                  45
Gln His Gly Lys Val Pro Leu Thr Phe Pro Pro Gly Gln Gly Val Pro
    50                  55                  60
Leu Asn Ala Asn Ser Thr Pro Ala Gln Asn Ala Gly Tyr Trp Arg Arg
65                  70                  75                  80
Gln Asp Arg Lys Ile Asn Thr Gly Asn Gly Ile Lys Gln Leu Ala Pro
                85                  90                  95
Arg Trp Tyr Phe Tyr Tyr Thr Gly Thr Gly Pro Glu Ala Ala Leu Pro
            100                 105                 110
Phe Arg Ala Val Lys Asp Gly Ile Val Trp Val His Glu Asp Gly Ala
        115                 120                 125
Thr Asp Ala Pro Ser Thr Phe Gly Thr Arg Asn Pro Asn Asn Asp Ser
    130                 135                 140
Ala Ile Val Thr Gln Phe Ala Pro Gly Thr Lys Leu Pro Lys Asn Phe
145                 150                 155                 160
His Ile Glu Gly Thr Gly Gly Asn Ser Gln Ser Ser Ser Arg Ala Ser
                165                 170                 175
Ser Leu Ser Arg Asn Ser Ser Arg Ser Ser Ser Gln Gly Ser Arg Ser
            180                 185                 190
Gly Asn Ser Thr Arg Gly Thr Ser Pro Gly Pro Ser Gly Ile Gly Ala
        195                 200                 205
Val Gly Gly Asp Leu Leu Tyr Leu Asp Leu Leu Asn Arg Leu Gln Ala
    210                 215                 220
Leu Glu Ser Gly Lys Val Lys Gln Ser Gln Pro Lys Val Ile Thr Lys
225                 230                 235                 240
Lys Asp Ala Ala Ala Ala Lys Asn Lys Met Arg His Lys Arg Thr Ser
                245                 250                 255
Thr Lys Ser Phe Asn Met Val Gln Ala Phe Gly Leu Arg Gly Pro Gly
            260                 265                 270
Asp Leu Gln Gly Asn Phe Gly Asp Leu Gln Leu Asn Lys Leu Gly Thr
        275                 280                 285
Glu Asp Pro Arg Trp Pro Gln Ile Ala Glu Leu Ala Pro Thr Ala Ser
    290                 295                 300
Ala Phe Met Gly Met Ser Gln Phe Lys Leu Thr His Gln Asn Asn Asp
305                 310                 315                 320
```

```
Asp His Gly Asn Pro Val Tyr Phe Leu Arg Tyr Ser Gly Ala Ile Lys
                325                 330                 335
Leu Asp Pro Lys Asn Pro Asn Tyr Asn Lys Trp Leu Glu Leu Leu Glu
            340                 345                 350
Gln Asn Ile Asp Ala Tyr Lys Thr Phe Pro Lys Lys Glu Lys Lys Gln
        355                 360                 365
Lys Ala Pro Lys Glu Glu Ser Thr Asp Gln Met Ser Glu Pro Pro Lys
    370                 375                 380
Glu Gln Arg Val Gln Gly Ser Ile Thr Gln Arg Thr Arg Thr Arg Pro
385                 390                 395                 400
Ser Val Gln Pro Gly Pro Met Ile Asp Val Asn Thr Asp
                405                 410


<210> 10
<211> 422
<212> PRT
<213> SARS-COV-1

<400> 10
Met Ser Asp Asn Gly Pro Gln Ser Asn Gln Arg Ser Ala Pro Arg Ile
 1               5                  10                  15
Thr Phe Gly Gly Pro Thr Asp Ser Thr Asp Asn Asn Gln Asn Gly Gly
            20                  25                  30
Arg Asn Gly Ala Arg Pro Lys Gln Arg Arg Pro Gln Gly Leu Pro Asn
        35                  40                  45
Asn Thr Ala Ser Trp Phe Thr Ala Leu Thr Gln His Gly Lys Glu Glu
    50                  55                  60
Leu Arg Phe Pro Arg Gly Gln Gly Val Pro Ile Asn Thr Asn Ser Gly
65                  70                  75                  80
Pro Asp Asp Gln Ile Gly Tyr Tyr Arg Arg Ala Thr Arg Arg Val Arg
                85                  90                  95
Gly Gly Asp Gly Lys Met Lys Glu Leu Ser Pro Arg Trp Tyr Phe Tyr
            100                 105                 110
Tyr Leu Gly Thr Gly Pro Glu Ala Ser Leu Pro Tyr Gly Ala Asn Lys
        115                 120                 125
Glu Gly Ile Val Trp Val Ala Thr Glu Gly Ala Leu Asn Thr Pro Lys
    130                 135                 140
Asp His Ile Gly Thr Arg Asn Pro Asn Asn Asn Ala Ala Thr Val Leu
145                 150                 155                 160
Gln Leu Pro Gln Gly Thr Thr Leu Pro Lys Gly Phe Tyr Ala Glu Gly
                165                 170                 175
Ser Arg Gly Gly Ser Gln Ala Ser Ser Arg Ser Ser Ser Arg Ser Arg
            180                 185                 190
Gly Asn Ser Arg Asn Ser Thr Pro Gly Ser Ser Arg Gly Asn Ser Pro
        195                 200                 205
Ala Arg Met Ala Ser Gly Gly Gly Glu Thr Ala Leu Ala Leu Leu Leu
    210                 215                 220
Leu Asp Arg Leu Asn Gln Leu Glu Ser Lys Val Ser Gly Lys Gly Gln
225                 230                 235                 240
```

```
Gln Gln Gln Gly Gln Thr Val Thr Lys Lys Ser Ala Ala Glu Ala Ser
                245                 250                 255
Lys Lys Pro Arg Gln Lys Arg Thr Ala Thr Lys Gln Tyr Asn Val Thr
            260                 265                 270
Gln Ala Phe Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly
        275                 280                 285
Asp Gln Asp Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln
    290                 295                 300
Ile Ala Gln Phe Ala Pro Ser Ala Ser Ala Phe Phe Gly Met Ser Arg
305                 310                 315                 320
Ile Gly Met Glu Val Thr Pro Ser Gly Thr Trp Leu Thr Tyr His Gly
                325                 330                 335
Ala Ile Lys Leu Asp Asp Lys Asp Pro Gln Phe Lys Asp Asn Val Ile
            340                 345                 350
Leu Leu Asn Lys His Ile Asp Ala Tyr Lys Thr Phe Pro Pro Thr Glu
        355                 360                 365
Pro Lys Lys Asp Lys Lys Lys Lys Thr Asp Glu Ala Gln Pro Leu Pro
    370                 375                 380
Gln Arg Gln Lys Lys Gln Pro Thr Val Thr Leu Leu Pro Ala Ala Asp
385                 390                 395                 400
Met Asp Asp Phe Ser Arg Gln Leu Gln Asn Ser Met Ser Gly Ala Ser
                405                 410                 415
Ala Asp Ser Thr Gln Ala
            420
```

<210> 11
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 11
ggcgcgccat tggccaccgc ggccgc 26

<210> 12
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 12
attggccacc 10

<210> 13
<211> 497

<212> PRT
<213> Influenza A Virus

<400> 13
Ala Ser Gln Gly Thr Lys Arg Ser Tyr Glu Gln Met Glu Thr Gly Gly
 1   5   10   15
Glu Arg Gln Asn Ala Thr Glu Ile Arg Ala Ser Val Gly Arg Met Val
            20   25   30
Gly Gly Ile Gly Arg Phe Tyr Ile Gln Met Cys Thr Glu Leu Lys Leu
        35   40   45
Ser Asp Tyr Glu Gly Arg Leu Ile Gln Asn Ser Ile Thr Ile Glu Arg
    50   55   60
Met Val Leu Ser Ala Phe Asp Glu Arg Arg Asn Lys Tyr Leu Glu Glu
65   70   75   80
His Pro Ser Ala Gly Lys Asp Pro Lys Lys Thr Gly Gly Pro Ile Tyr
                85   90   95
Arg Arg Arg Asp Gly Lys Trp Val Arg Glu Leu Ile Leu Tyr Asp Lys
            100   105   110
Glu Glu Ile Arg Arg Ile Trp Arg Gln Ala Asn Asn Gly Glu Asp Ala
        115   120   125
Thr Ala Gly Leu Thr His Leu Met Ile Trp His Ser Asn Leu Asn Asp
    130   135   140
Ala Thr Tyr Gln Arg Thr Arg Ala Leu Val Arg Thr Gly Met Asp Pro
145   150   155   160
Arg Met Cys Ser Leu Met Gln Gly Ser Thr Leu Pro Arg Arg Ser Gly
                165   170   175
Ala Ala Gly Ala Ala Val Lys Gly Val Gly Thr Met Val Met Glu Leu
            180   185   190
Ile Arg Met Ile Lys Arg Gly Ile Asn Asp Arg Asn Phe Trp Arg Gly
        195   200   205
Glu Asn Gly Arg Arg Thr Arg Ile Ala Tyr Glu Arg Met Cys Asn Ile
    210   215   220
Leu Lys Gly Lys Phe Gln Thr Ala Ala Gln Arg Ala Met Met Asp Gln
225   230   235   240
Val Arg Glu Ser Arg Asn Pro Gly Asn Ala Glu Ile Glu Asp Leu Ile
                245   250   255
Phe Leu Ala Arg Ser Ala Leu Ile Leu Arg Gly Ser Val Ala His Lys
            260   265   270
Ser Cys Leu Pro Ala Cys Val Tyr Gly Leu Ala Val Ala Ser Gly Tyr
        275   280   285
Asp Phe Glu Arg Glu Gly Tyr Ser Leu Val Gly Ile Asp Pro Phe Arg
    290   295   300
Leu Leu Gln Asn Ser Gln Val Phe Ser Leu Ile Arg Pro Asn Glu Asn
305   310   315   320
Pro Ala His Lys Ser Gln Leu Val Trp Met Ala Cys His Ser Ala Ala
                325   330   335
Phe Glu Asp Leu Arg Val Ser Ser Phe Ile Arg Gly Thr Arg Val Val
            340   345   350
Pro Arg Gly Gln Leu Ser Thr Arg Gly Val Gln Ile Ala Ser Asn Glu
        355   360   365

```
Asn Met Glu Thr Met Asp Ser Ser Thr Leu Glu Leu Arg Ser Arg Tyr
    370                 375                 380
Trp Ala Ile Arg Thr Arg Ser Gly Gly Thr Thr Asn Gln Gln Arg Ala
385                 390                 395                 400
Ser Ala Gly Gln Ile Ser Val Gln Pro Thr Phe Ser Val Gln Arg Asn
                405                 410                 415
Leu Pro Phe Glu Arg Ala Thr Ile Met Ala Ala Phe Thr Gly Asn Thr
            420                 425                 430
Glu Gly Arg Thr Ser Asp Met Arg Thr Glu Ile Ile Arg Met Met Glu
        435                 440                 445
Ser Ala Lys Pro Glu Asp Val Ser Phe Gln Gly Arg Gly Val Phe Glu
    450                 455                 460
Leu Ser Asp Glu Lys Ala Thr Asn Pro Ile Val Pro Ser Phe Asp Met
465                 470                 475                 480
Ser Asn Glu Gly Ser Tyr Phe Phe Gly Asp Asn Ala Glu Glu Tyr Asp
                485                 490                 495
Asn
```

<210> 14
<211> 560
<212> PRT
<213> Influenza B Virus

<400> 14

```
Met Ser Asn Met Asp Ile Asp Gly Ile Asn Thr Gly Thr Ile Asp Lys
 1               5                  10                  15
Thr Pro Glu Glu Ile Thr Pro Gly Thr Ser Gly Thr Thr Arg Pro Ile
            20                  25                  30
Ile Arg Pro Ala Thr Leu Ala Pro Pro Ser Asn Lys Arg Thr Arg Asn
        35                  40                  45
Pro Ser Pro Glu Arg Ala Thr Thr Ser Ser Glu Asp Asp Val Gly Arg
    50                  55                  60
Lys Thr Gln Lys Lys Gln Thr Pro Thr Glu Ile Lys Lys Ser Val Tyr
65                  70                  75                  80
Asn Met Val Val Lys Leu Gly Glu Phe Tyr Asn Gln Met Met Val Lys
                85                  90                  95
Ala Gly Leu Asn Asp Asp Met Glu Arg Asn Leu Ile Gln Asn Ala His
            100                 105                 110
Ala Val Glu Arg Ile Leu Leu Ala Ala Thr Asp Asp Lys Lys Thr Glu
        115                 120                 125
Phe Gln Lys Lys Lys Asn Ala Arg Asp Val Lys Glu Gly Lys Glu Glu
    130                 135                 140
Ile Asp His Asn Lys Thr Gly Gly Thr Phe Tyr Lys Met Val Arg Asp
145                 150                 155                 160
Asp Lys Thr Ile Tyr Phe Ser Pro Ile Arg Ile Thr Phe Leu Lys Glu
                165                 170                 175
Glu Val Lys Thr Met Tyr Lys Thr Thr Met Gly Ser Asp Gly Phe Ser
            180                 185                 190
```

```
Gly Leu Asn His Ile Met Ile Gly His Ser Gln Met Asn Asp Val Cys
        195                 200                 205
Phe Gln Arg Ser Lys Ala Leu Lys Arg Val Gly Leu Asp Pro Ser Leu
    210                 215                 220
Ile Ser Thr Phe Ala Gly Ser Thr Ile Pro Arg Arg Ser Gly Ala Thr
225                 230                 235                 240
Gly Val Ala Ile Lys Gly Gly Gly Thr Leu Val Ala Glu Ala Ile Arg
                245                 250                 255
Phe Ile Gly Arg Ala Met Ala Asp Arg Gly Leu Leu Arg Asp Ile Lys
            260                 265                 270
Ala Lys Thr Ala Tyr Glu Lys Ile Leu Leu Asn Leu Lys Asn Lys Cys
        275                 280                 285
Ser Ala Pro Gln Gln Lys Ala Leu Val Asp Gln Val Ile Gly Ser Arg
    290                 295                 300
Asn Pro Gly Ile Ala Asp Ile Glu Asp Leu Thr Leu Leu Ala Arg Ser
305                 310                 315                 320
Met Val Val Val Arg Pro Ser Val Ala Ser Lys Val Val Leu Pro Ile
                325                 330                 335
Ser Ile Tyr Ala Lys Ile Pro Gln Leu Gly Phe Asn Val Glu Glu Tyr
            340                 345                 350
Ser Met Val Gly Tyr Glu Ala Met Ala Leu Tyr Asn Met Ala Thr Pro
        355                 360                 365
Val Ser Ile Leu Arg Met Gly Asp Asp Ala Lys Asp Lys Ser Gln Leu
    370                 375                 380
Phe Phe Met Ser Cys Phe Gly Ala Ala Tyr Glu Asp Leu Arg Val Leu
385                 390                 395                 400
Ser Ala Leu Thr Gly Thr Glu Phe Lys Pro Arg Ser Ala Leu Lys Cys
                405                 410                 415
Lys Gly Phe His Val Pro Ala Lys Glu Gln Val Glu Gly Met Gly Ala
            420                 425                 430
Ala Leu Met Ser Ile Lys Leu Gln Phe Trp Ala Pro Met Thr Arg Ser
        435                 440                 445
Gly Gly Asn Glu Val Gly Gly Asp Gly Gly Ser Gly Gln Ile Ser Cys
    450                 455                 460
Ser Pro Val Phe Ala Val Glu Arg Pro Ile Ala Leu Ser Lys Gln Ala
465                 470                 475                 480
Val Arg Arg Met Leu Ser Met Asn Ile Glu Gly Arg Asp Ala Asp Val
                485                 490                 495
Lys Gly Asn Leu Leu Lys Met Met Asn Asp Ser Met Ala Lys Lys Thr
            500                 505                 510
Ser Gly Asn Ala Phe Ile Gly Lys Lys Met Phe Gln Ile Ser Asp Lys
        515                 520                 525
Asn Lys Thr Asn Pro Val Glu Ile Pro Ile Lys Gln Thr Ile Pro Asn
    530                 535                 540
Phe Phe Phe Gly Arg Asp Thr Ala Glu Asp Tyr Asp Asp Leu Asp Tyr
545                 550                 555                 560
```

<210> 15
<211> 3246

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 15
atgcctatgg gcagcctgca gccactggct acactgtacc tgctgggcat gctggtggcc 60
tcttgtctgg gcgccagcca aggcactaag agaagctacg agcagatgga aaccggaggc 120
gaacggcaga acgccacaga gatcagagcc tctgtgggcc gtatggtcgg cggcatcggc 180
agattctaca tccagatgtg caccgaactg aagctgagcg actacgaggg ccgcctgatc 240
cagaacagca tcacaatcga gagaatggtg ctgtccgcct ttgacgagcg gagaaacaaa 300
tacctggaag agcaccctag cgccggaaaa gatcctaaga aaaccggcgg acctatctac 360
agaagaagag atggtaagtg ggtgagagag ctgattctgt acgataagga agagattcga 420
agaatctgga gacaggccaa caacggcgag gatgccaccg caggcctgac acacctgatg 480
atctggcaca gcaacctgaa cgatgcgacc taccagcgca cgcgggccct ggtcagaacc 540
ggcatggatc ctcggatgtg tagcctgatg cagggcagca cactgccaag acggagtggg 600
gccgccggcg ctgcagtgaa gggcgtcgga accatggtga tggagctgat ccggatgata 660
aagcggggca tcaacgacag aaacttctgg cgaggcgaga acggccgaag aacccggatc 720
gcctacgaga gaatgtgcaa catcctgaaa ggaaaattcc agaccgccgc ccagcgggcc 780
atgatggacc aggtgcgcga gagcagaaac cccggcaatg ccgagatcga ggacctgatc 840
ttcctggcca gaagcgccct cattcttaga ggctctgtgg cccacaagag ctgtctgcct 900
gcctgtgtgt acggcctggc agtggcctca ggctacgact tcgagcggga aggatacagt 960
ctggtgggca tcgacccttt cagactcctg cagaatagcc aggtgtttag cctgatcaga 1020
ccaaacgaaa accccgccca taagagccag ctggtgtgga tggcctgcca cagcgccgcc 1080
tttgaggatc tgagagtgag ctcttttatc agaggcaccc gggtggttcc acgaggtcaa 1140
ctgtctacaa gaggtgtgca gatcgccagc aacgagaaca tggagaccat ggatagcagc 1200
accctggaac tgagatccag atactgggcc atcaggacac ggagcggcgg caccaccaat 1260
cagcagcgcg ccagcgccgg ccagatctct gtccagccta cgtttagcgt gcagcggaat 1320
ttgcccttcg aacgcgccac aatcatggct gctttcaccg gcaatacaga gggcagaacc 1380
agcgatatga gaacagaaat tatccgtatg atggagtccg caaaacctga ggacgtgtcc 1440
ttccaaggca gaggcgtgtt cgagctgagc gacgagaagg ccaccaaccc tatcgtgcct 1500
agcttcgata tgtctaatga gggcagctac tttttcggag ataacgccga agagtacgac 1560
aacatgtcta atatggatat cgacggcatt aacaccggca ccatcgacaa aacccctgag 1620
gagatcaccc ctggcaccag cggcacaacc cggcccatca tccgccccgc tacactggct 1680
ccacctagca acaagcggac cagaaatccc tcgccagaaa gagccacaac ctccagcgag 1740
gacgacgtgg gacggaagac acaaaagaag cagacccta cagagatcaa gaagtctgtt 1800
tacaacatgg tggtgaaact gggcgagttc tacaaccaga tgatggtgaa ggccggcctg 1860
aacgacgata tggaaagaaa tctgatccag aacgcccacg ccgtggagcg gattctgctg 1920
gccgccaccg atgataagaa gaccgaattc cagaaaaaga aaaacgccag agacgtgaag 1980
gaaggcaagg aagagatcga ccacaacaag acaggcggca cattctacaa gatggtccgg 2040
gacgacaaga ccatctactt cagccctatc cggataacat tcctgaaaga agaagtgaag 2100
accatgtaca aaaccacaat gggctctgac ggcttcagcg gcctgaatca catcatgatc 2160
ggccactctc aaatgaacga tgtgtgcttc cagagaagca aggctctgaa gcgcgtgggc 2220
ctggatccta gcctgatctc taccttcgcc ggcagcacca tccccagaag atcgggcgct 2280
accggcgtgg ctatcaaggg aggaggcaca ctggtggctg aagccatcag attcatcgga 2340
agagccatgg ccgacagagg actcctgaga gatatcaaag ccaaaaccgc ctacgaaaaa 2400
atcctgctga acctgaagaa caagtgcagc gcgcctcaac agaaggccct ggtggaccag 2460
gttatcggct ctagaaaccc tggaatcgcc gatatcgagg acctgacact gctggccaga 2520
tctatggtgg tggtgagacc ctccgtggcc agcaaggtgg tgctgcctat cagcatctac 2580

gccaagatcc ctcagctggg atttaacgtg gaagaataca gcatggttgg ttatgaggcc 2640
atggccctgt acaacatggc cacacctgtg tccatcctga gaatgggcga cgatgccaaa 2700
gacaagagcc agctgttctt catgagctgc ttcggcgctg cctatgagga cctgagagtg 2760
ctgtccgctc ttacaggaac agagttcaag cctaggagcg cactgaagtg caagggcttc 2820
cacgtgcccg ccaaggaaca ggtggaaggc atgggagctg ctctgatgtc catcaagctg 2880
caattttggg ctcctatgac ccggagcggc ggaaatgagg tgggtggcga cggaggcagc 2940
ggacagattt cttgcagccc cgtatttgcc gtggagagac caatcgccct gtccaagcag 3000
gccgtgagaa gaatgctgag catgaacatc gagggccggg acgccgacgt gaagggcaac 3060
ctgttgaaga tgatgaacga cagcatggcc aagaagacca gtggcaatgc cttcatcggc 3120
aagaagatgt tccagatctc cgacaagaac aagaccaacc ccgtggaaat ccccatcaag 3180
cagacaatcc ctaacttctt cttcggcaga gacaccgccg aagactatga cgacctggac 3240
tactga 3246

<210> 16
<211> 1081
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 16
Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1   5   10   15
Met Leu Val Ala Ser Cys Leu Gly Ala Ser Gln Gly Thr Lys Arg Ser
            20   25   30
Tyr Glu Gln Met Glu Thr Gly Gly Glu Arg Gln Asn Ala Thr Glu Ile
        35   40   45
Arg Ala Ser Val Gly Arg Met Val Gly Gly Ile Gly Arg Phe Tyr Ile
    50   55   60
Gln Met Cys Thr Glu Leu Lys Leu Ser Asp Tyr Glu Gly Arg Leu Ile
65   70   75   80
Gln Asn Ser Ile Thr Ile Glu Arg Met Val Leu Ser Ala Phe Asp Glu
                85   90   95
Arg Arg Asn Lys Tyr Leu Glu Glu His Pro Ser Ala Gly Lys Asp Pro
            100   105   110
Lys Lys Thr Gly Gly Pro Ile Tyr Arg Arg Arg Asp Gly Lys Trp Val
        115   120   125
Arg Glu Leu Ile Leu Tyr Asp Lys Glu Glu Ile Arg Arg Ile Trp Arg
    130   135   140
Gln Ala Asn Asn Gly Glu Asp Ala Thr Ala Gly Leu Thr His Leu Met
145   150   155   160
Ile Trp His Ser Asn Leu Asn Asp Ala Thr Tyr Gln Arg Thr Arg Ala
                165   170   175
Leu Val Arg Thr Gly Met Asp Pro Arg Met Cys Ser Leu Met Gln Gly
            180   185   190
Ser Thr Leu Pro Arg Arg Ser Gly Ala Ala Gly Ala Ala Val Lys Gly
        195   200   205
Val Gly Thr Met Val Met Glu Leu Ile Arg Met Ile Lys Arg Gly Ile
    210   215   220

```
Asn Asp Arg Asn Phe Trp Arg Gly Glu Asn Gly Arg Arg Thr Arg Ile
225                 230                 235                 240
Ala Tyr Glu Arg Met Cys Asn Ile Leu Lys Gly Lys Phe Gln Thr Ala
                245                 250                 255
Ala Gln Arg Ala Met Met Asp Gln Val Arg Glu Ser Arg Asn Pro Gly
            260                 265                 270
Asn Ala Glu Ile Glu Asp Leu Ile Phe Leu Ala Arg Ser Ala Leu Ile
        275                 280                 285
Leu Arg Gly Ser Val Ala His Lys Ser Cys Leu Pro Ala Cys Val Tyr
    290                 295                 300
Gly Leu Ala Val Ala Ser Gly Tyr Asp Phe Glu Arg Glu Gly Tyr Ser
305                 310                 315                 320
Leu Val Gly Ile Asp Pro Phe Arg Leu Leu Gln Asn Ser Gln Val Phe
                325                 330                 335
Ser Leu Ile Arg Pro Asn Glu Asn Pro Ala His Lys Ser Gln Leu Val
            340                 345                 350
Trp Met Ala Cys His Ser Ala Ala Phe Glu Asp Leu Arg Val Ser Ser
        355                 360                 365
Phe Ile Arg Gly Thr Arg Val Val Pro Arg Gly Gln Leu Ser Thr Arg
    370                 375                 380
Gly Val Gln Ile Ala Ser Asn Glu Asn Met Glu Thr Met Asp Ser Ser
385                 390                 395                 400
Thr Leu Glu Leu Arg Ser Arg Tyr Trp Ala Ile Arg Thr Arg Ser Gly
                405                 410                 415
Gly Thr Thr Asn Gln Gln Arg Ala Ser Ala Gly Gln Ile Ser Val Gln
            420                 425                 430
Pro Thr Phe Ser Val Gln Arg Asn Leu Pro Phe Glu Arg Ala Thr Ile
        435                 440                 445
Met Ala Ala Phe Thr Gly Asn Thr Glu Gly Arg Thr Ser Asp Met Arg
    450                 455                 460
Thr Glu Ile Ile Arg Met Met Glu Ser Ala Lys Pro Glu Asp Val Ser
465                 470                 475                 480
Phe Gln Gly Arg Gly Val Phe Glu Leu Ser Asp Glu Lys Ala Thr Asn
                485                 490                 495
Pro Ile Val Pro Ser Phe Asp Met Ser Asn Glu Gly Ser Tyr Phe Phe
            500                 505                 510
Gly Asp Asn Ala Glu Glu Tyr Asp Asn Met Ser Asn Met Asp Ile Asp
        515                 520                 525
Gly Ile Asn Thr Gly Thr Ile Asp Lys Thr Pro Glu Glu Ile Thr Pro
    530                 535                 540
Gly Thr Ser Gly Thr Thr Arg Pro Ile Ile Arg Pro Ala Thr Leu Ala
545                 550                 555                 560
Pro Pro Ser Asn Lys Arg Thr Arg Asn Pro Ser Pro Glu Arg Ala Thr
                565                 570                 575
Thr Ser Ser Glu Asp Asp Val Gly Arg Lys Thr Gln Lys Lys Gln Thr
            580                 585                 590
Pro Thr Glu Ile Lys Lys Ser Val Tyr Asn Met Val Val Lys Leu Gly
        595                 600                 605
Glu Phe Tyr Asn Gln Met Met Val Lys Ala Gly Leu Asn Asp Asp Met
    610                 615                 620
```

```
Glu Arg Asn Leu Ile Gln Asn Ala His Ala Val Glu Arg Ile Leu Leu
625                 630                 635                 640
Ala Ala Thr Asp Asp Lys Lys Thr Glu Phe Gln Lys Lys Lys Asn Ala
                645                 650                 655
Arg Asp Val Lys Glu Gly Lys Glu Glu Ile Asp His Asn Lys Thr Gly
            660                 665                 670
Gly Thr Phe Tyr Lys Met Val Arg Asp Asp Lys Thr Ile Tyr Phe Ser
        675                 680                 685
Pro Ile Arg Ile Thr Phe Leu Lys Glu Glu Val Lys Thr Met Tyr Lys
    690                 695                 700
Thr Thr Met Gly Ser Asp Gly Phe Ser Gly Leu Asn His Ile Met Ile
705                 710                 715                 720
Gly His Ser Gln Met Asn Asp Val Cys Phe Gln Arg Ser Lys Ala Leu
                725                 730                 735
Lys Arg Val Gly Leu Asp Pro Ser Leu Ile Ser Thr Phe Ala Gly Ser
            740                 745                 750
Thr Ile Pro Arg Arg Ser Gly Ala Thr Gly Val Ala Ile Lys Gly Gly
        755                 760                 765
Gly Thr Leu Val Ala Glu Ala Ile Arg Phe Ile Gly Arg Ala Met Ala
    770                 775                 780
Asp Arg Gly Leu Leu Arg Asp Ile Lys Ala Lys Thr Ala Tyr Glu Lys
785                 790                 795                 800
Ile Leu Leu Asn Leu Lys Asn Lys Cys Ser Ala Pro Gln Gln Lys Ala
                805                 810                 815
Leu Val Asp Gln Val Ile Gly Ser Arg Asn Pro Gly Ile Ala Asp Ile
            820                 825                 830
Glu Asp Leu Thr Leu Leu Ala Arg Ser Met Val Val Val Arg Pro Ser
        835                 840                 845
Val Ala Ser Lys Val Val Leu Pro Ile Ser Ile Tyr Ala Lys Ile Pro
    850                 855                 860
Gln Leu Gly Phe Asn Val Glu Glu Tyr Ser Met Val Gly Tyr Glu Ala
865                 870                 875                 880
Met Ala Leu Tyr Asn Met Ala Thr Pro Val Ser Ile Leu Arg Met Gly
                885                 890                 895
Asp Asp Ala Lys Asp Lys Ser Gln Leu Phe Phe Met Ser Cys Phe Gly
            900                 905                 910
Ala Ala Tyr Glu Asp Leu Arg Val Leu Ser Ala Leu Thr Gly Thr Glu
        915                 920                 925
Phe Lys Pro Arg Ser Ala Leu Lys Cys Lys Gly Phe His Val Pro Ala
    930                 935                 940
Lys Glu Gln Val Glu Gly Met Gly Ala Ala Leu Met Ser Ile Lys Leu
945                 950                 955                 960
Gln Phe Trp Ala Pro Met Thr Arg Ser Gly Gly Asn Glu Val Gly Gly
                965                 970                 975
Asp Gly Gly Ser Gly Gln Ile Ser Cys Ser Pro Val Phe Ala Val Glu
            980                 985                 990
Arg Pro Ile Ala Leu Ser Lys Gln Ala Val Arg Arg Met Leu Ser Met
        995                 1000                1005
Asn Ile Glu Gly Arg Asp Ala Asp Val Lys Gly Asn Leu Leu Lys Met
    1010                1015                1020
```

```
Met Asn Asp Ser Met Ala Lys Lys Thr Ser Gly Asn Ala Phe Ile Gly
1025                1030                1035                1040
Lys Lys Met Phe Gln Ile Ser Asp Lys Asn Lys Thr Asn Pro Val Glu
                1045                1050                1055
Ile Pro Ile Lys Gln Thr Ile Pro Asn Phe Phe Phe Gly Arg Asp Thr
            1060                1065                1070
Ala Glu Asp Tyr Asp Asp Leu Asp Tyr
        1075                1080


<210> 17
<211> 497
<212> PRT
<213> Influenza A Virus

<400> 17
Ala Ser Gln Gly Thr Lys Arg Ser Tyr Glu Gln Met Glu Thr Asp Gly
 1               5                  10                  15
Glu Arg Gln Asn Ala Asn Glu Ile Arg Ala Ser Val Gly Lys Met Ile
            20                  25                  30
Gly Gly Ile Gly Arg Phe Tyr Ile Gln Met Cys Thr Glu Leu Lys Leu
        35                  40                  45
Ser Asp Tyr Glu Gly Arg Leu Ile Gln Asn Ser Leu Thr Ile Glu Arg
    50                  55                  60
Met Val Leu Ser Ala Phe Asp Glu Arg Arg Asn Lys Tyr Leu Glu Glu
65                  70                  75                  80
His Pro Ser Ala Gly Lys Asp Pro Lys Lys Thr Gly Gly Pro Ile Tyr
                85                  90                  95
Lys Arg Val Asp Gly Lys Trp Met Arg Glu Leu Val Leu Tyr Asp Lys
            100                 105                 110
Glu Glu Ile Arg Arg Ile Trp Arg Gln Ala Asn Asn Gly Asp Asp Ala
        115                 120                 125
Thr Ala Gly Leu Thr His Met Met Ile Trp His Ser Asn Leu Asn Asp
    130                 135                 140
Thr Thr Tyr Gln Arg Thr Arg Ala Leu Val Arg Thr Gly Met Asp Pro
145                 150                 155                 160
Arg Met Cys Ser Leu Met Gln Gly Ser Thr Leu Pro Arg Arg Ser Gly
                165                 170                 175
Ala Ala Gly Ala Ala Val Lys Gly Val Gly Thr Met Val Met Glu Leu
            180                 185                 190
Ile Arg Met Ile Lys Arg Gly Ile Asn Asp Arg Asn Phe Trp Arg Gly
        195                 200                 205
Glu Asn Gly Arg Lys Thr Arg Asn Ala Tyr Glu Arg Met Cys Asn Ile
    210                 215                 220
Leu Lys Gly Lys Phe Gln Thr Ala Ala Gln Arg Ala Met Met Asp Gln
225                 230                 235                 240
Val Arg Glu Ser Arg Asn Pro Gly Asn Ala Glu Ile Glu Asp Leu Ile
                245                 250                 255
Phe Leu Ala Arg Ser Ala Leu Ile Leu Arg Gly Ser Val Ala His Lys
            260                 265                 270
```

```
Ser Cys Leu Pro Ala Cys Val Tyr Gly Pro Ala Val Ala Ser Gly Tyr
        275                 280                 285
Asp Phe Glu Lys Glu Gly Tyr Ser Leu Val Gly Ile Asp Pro Phe Lys
    290                 295                 300
Leu Leu Gln Asn Ser Gln Val Tyr Ser Leu Ile Arg Pro Asn Glu Asn
305                 310                 315                 320
Pro Ala His Lys Ser Gln Leu Val Trp Met Ala Cys Asn Ser Ala Ala
                325                 330                 335
Phe Glu Asp Leu Arg Val Ser Ser Phe Ile Arg Gly Thr Lys Val Ile
            340                 345                 350
Pro Arg Gly Lys Leu Ser Thr Arg Gly Val Gln Ile Ala Ser Asn Glu
        355                 360                 365
Asn Met Asp Thr Met Gly Ser Ser Thr Leu Glu Leu Arg Ser Arg Tyr
    370                 375                 380
Trp Ala Ile Arg Thr Arg Ser Gly Gly Asn Thr Asn Gln Gln Arg Ala
385                 390                 395                 400
Ser Ala Gly Gln Ile Ser Val Gln Pro Thr Phe Ser Val Gln Arg Asn
                405                 410                 415
Leu Pro Phe Asp Lys Pro Thr Ile Met Ala Ala Phe Thr Gly Asn Ala
            420                 425                 430
Glu Gly Arg Thr Ser Asp Met Arg Ala Glu Ile Ile Arg Met Met Glu
        435                 440                 445
Gly Ala Lys Pro Glu Glu Val Ser Phe Gln Gly Arg Gly Val Phe Glu
    450                 455                 460
Leu Ser Asp Glu Lys Ala Thr Asn Pro Ile Val Pro Ser Phe Asp Met
465                 470                 475                 480
Ser Asn Glu Gly Ser Tyr Phe Phe Gly Asp Asn Ala Glu Glu Tyr Asp
                485                 490                 495
Asn
```

<210> 18
<211> 738
<212> PRT
<213> Zaire ebolavirus

<400> 18

```
Asp Ser Arg Pro Gln Lys Ile Trp Met Ala Pro Ser Leu Thr Glu Ser
 1               5                  10                  15
Asp Met Asp Tyr His Lys Ile Leu Thr Ala Gly Leu Ser Val Gln Gln
            20                  25                  30
Gly Ile Val Arg Gln Arg Val Ile Pro Val Tyr Gln Val Asn Asn Leu
        35                  40                  45
Glu Glu Ile Cys Gln Leu Ile Ile Gln Ala Phe Glu Ala Gly Val Asp
    50                  55                  60
Phe Gln Glu Ser Ala Asp Ser Phe Leu Leu Met Leu Cys Leu His His
65                  70                  75                  80
Ala Tyr Gln Gly Asp Tyr Lys Leu Phe Leu Glu Ser Gly Ala Val Lys
                85                  90                  95
```

```
Tyr Leu Glu Gly His Gly Phe Arg Phe Glu Val Lys Lys Arg Asp Gly
            100                 105                 110
Val Lys Arg Leu Glu Glu Leu Leu Pro Ala Val Ser Ser Gly Lys Asn
        115                 120                 125
Ile Lys Arg Thr Leu Ala Ala Met Pro Glu Glu Glu Thr Thr Glu Ala
    130                 135                 140
Asn Ala Gly Gln Phe Leu Ser Phe Ala Ser Leu Phe Leu Pro Lys Leu
145                 150                 155                 160
Val Val Gly Glu Lys Ala Cys Leu Glu Lys Val Gln Arg Gln Ile Gln
                165                 170                 175
Val His Ala Glu Gln Gly Leu Ile Gln Tyr Pro Thr Ala Trp Gln Ser
            180                 185                 190
Val Gly His Met Met Val Ile Phe Arg Leu Met Arg Thr Asn Phe Leu
        195                 200                 205
Ile Lys Phe Leu Leu Ile His Gln Gly Met His Met Val Ala Gly His
    210                 215                 220
Asp Ala Asn Asp Ala Val Ile Ser Asn Ser Val Ala Gln Ala Arg Phe
225                 230                 235                 240
Ser Gly Leu Leu Ile Val Lys Thr Val Leu Asp His Ile Leu Gln Lys
                245                 250                 255
Thr Glu Arg Gly Val Arg Leu His Pro Leu Ala Arg Thr Ala Lys Val
            260                 265                 270
Lys Asn Glu Val Asn Ser Phe Lys Ala Ala Leu Ser Ser Leu Ala Lys
        275                 280                 285
His Gly Glu Tyr Ala Pro Phe Ala Arg Leu Leu Asn Leu Ser Gly Val
    290                 295                 300
Asn Asn Leu Glu His Gly Leu Phe Pro Gln Leu Ser Ala Ile Ala Leu
305                 310                 315                 320
Gly Val Ala Thr Ala His Gly Ser Thr Leu Ala Gly Val Asn Val Gly
                325                 330                 335
Glu Gln Tyr Gln Gln Leu Arg Glu Ala Ala Thr Glu Ala Glu Lys Gln
            340                 345                 350
Leu Gln Gln Tyr Ala Glu Ser Arg Glu Leu Asp His Leu Gly Leu Asp
        355                 360                 365
Asp Gln Glu Lys Lys Ile Leu Met Asn Phe His Gln Lys Lys Asn Glu
    370                 375                 380
Ile Ser Phe Gln Gln Thr Asn Ala Met Val Thr Leu Arg Lys Glu Arg
385                 390                 395                 400
Leu Ala Lys Leu Thr Glu Ala Ile Thr Ala Ala Ser Leu Pro Lys Thr
                405                 410                 415
Ser Gly His Tyr Asp Asp Asp Asp Asp Ile Pro Phe Pro Gly Pro Ile
            420                 425                 430
Asn Asp Asp Asp Asn Pro Gly His Gln Asp Asp Asp Pro Thr Asp Ser
        435                 440                 445
Gln Asp Thr Thr Ile Pro Asp Val Val Val Asp Pro Asp Asp Gly Ser
    450                 455                 460
Tyr Gly Glu Tyr Gln Ser Tyr Ser Glu Asn Gly Met Asn Ala Pro Asp
465                 470                 475                 480
Asp Leu Val Leu Phe Asp Leu Asp Glu Asp Asp Glu Asp Thr Lys Pro
                485                 490                 495
```

```
Val Pro Asn Arg Ser Thr Lys Gly Gly Gln Gln Lys Asn Ser Gln Lys
            500                 505                 510
Gly Gln His Ile Glu Gly Arg Gln Thr Gln Phe Arg Pro Ile Gln Asn
        515                 520                 525
Val Pro Gly Pro His Arg Thr Ile His His Ala Ser Ala Pro Leu Thr
    530                 535                 540
Asp Asn Asp Arg Arg Asn Glu Pro Ser Gly Ser Thr Ser Pro Arg Met
545                 550                 555                 560
Leu Thr Pro Ile Asn Glu Glu Ala Asp Pro Leu Asp Asp Ala Asp Asp
                565                 570                 575
Glu Thr Ser Ser Leu Pro Pro Leu Glu Ser Asp Asp Glu Glu Gln Asp
            580                 585                 590
Arg Asp Gly Thr Ser Asn Arg Thr Pro Thr Val Ala Pro Pro Ala Pro
        595                 600                 605
Val Tyr Arg Asp His Ser Glu Lys Lys Glu Leu Pro Gln Asp Glu Gln
    610                 615                 620
Gln Asp Gln Asp His Thr Gln Glu Ala Arg Asn Gln Asp Ser Asp Asn
625                 630                 635                 640
Thr Gln Ser Glu His Ser Leu Glu Glu Met Tyr Arg His Ile Leu Arg
                645                 650                 655
Ser Gln Gly Pro Phe Asp Ala Val Leu Tyr Tyr His Met Met Lys Asp
            660                 665                 670
Glu Pro Val Val Phe Ser Thr Ser Asp Gly Lys Glu Tyr Thr Tyr Pro
        675                 680                 685
Asp Ser Leu Glu Glu Glu Tyr Pro Pro Trp Leu Thr Glu Lys Glu Ala
    690                 695                 700
Met Asn Glu Glu Asn Arg Phe Val Thr Leu Asp Gly Gln Gln Phe Tyr
705                 710                 715                 720
Trp Pro Val Met Asn His Lys Asn Lys Phe Met Ala Ile Leu Gln His
                725                 730                 735
His Gln
```

<210> 19
<211> 336
<212> PRT
<213> Sudan ebolavirus

<400> 19

```
Ala Lys Leu Thr Glu Ala Ile Thr Thr Ala Ser Lys Ile Lys Val Gly
 1               5                  10                  15
Asp Arg Tyr Pro Asp Asp Asn Asp Ile Pro Phe Pro Gly Pro Ile Tyr
            20                  25                  30
Asp Asp Thr His Pro Asn Pro Ser Asp Asp Asn Pro Asp Asp Ser Arg
        35                  40                  45
Asp Thr Thr Ile Pro Gly Gly Val Val Asp Pro Tyr Asp Asp Glu Ser
    50                  55                  60
Asn Asn Tyr Pro Asp Tyr Glu Asp Ser Ala Glu Gly Thr Thr Gly Asp
65                  70                  75                  80
```

```
Leu Asp Leu Phe Asn Leu Asp Asp Asp Asp Asp Asp Ser Arg Pro Gly
                85                  90                  95
Pro Pro Asp Arg Gly Gln Asn Lys Glu Arg Ala Ala Arg Thr Tyr Gly
            100                 105                 110
Leu Gln Asp Pro Thr Leu Asp Gly Ala Lys Lys Val Pro Glu Leu Thr
        115                 120                 125
Pro Gly Ser His Gln Pro Gly Asn Leu His Ile Thr Lys Ser Gly Ser
    130                 135                 140
Asn Thr Asn Gln Pro Gln Gly Asn Met Ser Ser Thr Leu His Ser Met
145                 150                 155                 160
Thr Pro Ile Gln Glu Glu Ser Glu Pro Asp Asp Gln Lys Asp Asn Asp
                165                 170                 175
Asp Glu Ser Leu Thr Ser Leu Asp Ser Glu Gly Asp Glu Asp Gly Glu
            180                 185                 190
Ser Ile Ser Glu Glu Asn Thr Pro Thr Val Ala Pro Pro Ala Pro Val
        195                 200                 205
Tyr Lys Asp Thr Gly Val Asp Thr Asn Gln Gln Asn Gly Pro Ser Ser
    210                 215                 220
Thr Val Asp Ser Gln Gly Ser Glu Ser Glu Ala Leu Pro Ile Asn Ser
225                 230                 235                 240
Lys Lys Ser Ser Ala Leu Glu Glu Thr Tyr Tyr His Leu Leu Lys Thr
                245                 250                 255
Gln Gly Pro Phe Glu Ala Ile Asn Tyr Tyr His Leu Met Ser Asp Glu
            260                 265                 270
Pro Ile Ala Phe Ser Thr Glu Ser Gly Lys Glu Tyr Ile Phe Pro Asp
        275                 280                 285
Ser Leu Glu Glu Ala Tyr Pro Pro Trp Leu Ser Glu Lys Glu Ala Leu
    290                 295                 300
Glu Lys Glu Asn Arg Tyr Leu Val Ile Asp Gly Gln Gln Phe Leu Trp
305                 310                 315                 320
Pro Val Met Ser Leu Arg Asp Lys Phe Leu Ala Val Leu Gln His Asp
                325                 330                 335
```

<210> 20
<211> 337
<212> PRT
<213> Bundibugyo ebolavirus

<400> 20

```
Ala Lys Leu Thr Glu Ala Ile Thr Ser Thr Ser Ile Leu Lys Thr Gly
 1               5                  10                  15
Arg Arg Tyr Asp Asp Asp Asn Asp Ile Pro Phe Pro Gly Pro Ile Asn
            20                  25                  30
Asp Asn Glu Asn Ser Gly Gln Asn Asp Asp Asp Pro Thr Asp Ser Gln
        35                  40                  45
Asp Thr Thr Ile Pro Asp Val Ile Ile Asp Pro Asn Asp Gly Gly Tyr
    50                  55                  60
Asn Asn Tyr Ser Asp Tyr Ala Asn Asp Ala Ala Ser Ala Pro Asp Asp
65                  70                  75                  80
```

```
Leu Val Leu Phe Asp Leu Glu Asp Glu Asp Asp Ala Asp Asn Pro Ala
                85                  90                  95
Gln Asn Thr Pro Glu Lys Asn Asp Arg Pro Ala Thr Thr Lys Leu Arg
            100                 105                 110
Asn Gly Gln Asp Gln Asp Gly Asn Gln Gly Glu Thr Ala Ser Pro Arg
        115                 120                 125
Val Ala Pro Asn Gln Tyr Arg Asp Lys Pro Met Pro Gln Val Gln Asp
    130                 135                 140
Arg Ser Glu Asn His Asp Gln Thr Leu Gln Thr Gln Ser Arg Val Leu
145                 150                 155                 160
Thr Pro Ile Ser Glu Glu Ala Asp Pro Ser Asp His Asn Asp Gly Asp
                165                 170                 175
Asn Glu Ser Ile Pro Pro Leu Glu Ser Asp Asp Glu Gly Ser Thr Asp
            180                 185                 190
Thr Thr Ala Ala Glu Thr Lys Pro Ala Thr Ala Pro Pro Ala Pro Val
        195                 200                 205
Tyr Arg Ser Ile Ser Val Asp Asp Ser Val Pro Ser Glu Asn Ile Pro
    210                 215                 220
Ala Gln Ser Asn Gln Thr Asn Asn Glu Asp Asn Val Arg Asn Asn Ala
225                 230                 235                 240
Gln Ser Glu Gln Ser Ile Ala Glu Met Tyr Gln His Ile Leu Lys Thr
                245                 250                 255
Gln Gly Pro Phe Asp Ala Ile Leu Tyr Tyr His Met Met Lys Glu Glu
            260                 265                 270
Pro Ile Ile Phe Ser Thr Ser Asp Gly Lys Glu Tyr Thr Tyr Pro Asp
        275                 280                 285
Ser Leu Glu Asp Glu Tyr Pro Pro Trp Leu Ser Glu Lys Glu Ala Met
    290                 295                 300
Asn Glu Asp Asn Arg Phe Ile Thr Met Asp Gly Gln Gln Phe Tyr Trp
305                 310                 315                 320
Pro Val Met Asn His Arg Asn Lys Phe Met Ala Ile Leu Gln His His
                325                 330                 335
Arg
```

<210> 21
<211> 169
<212> PRT
<213> Tai Forest ebolavirus

<400> 21

```
Leu Val Leu Phe Asp Leu Glu Asp Gly Asp Glu Asp Asp His Arg Pro
 1               5                  10                  15
Ser Ser Ser Ser Glu Asn Asn Asn Lys His Ser Leu Thr Gly Thr Asp
            20                  25                  30
Ser Asn Lys Thr Ser Asn Trp Asn Arg Asn Pro Thr Asn Met Pro Lys
        35                  40                  45
Lys Asp Ser Thr Gln Asn Asn Asp Asn Pro Ala Gln Arg Ala Gln Glu
    50                  55                  60
```

```
Tyr Ala Arg Asp Asn Ile Gln Asp Thr Pro Thr Pro His Arg Ala Leu
65                  70                  75                  80
Thr Pro Ile Ser Glu Glu Thr Gly Ser Asn Gly His Asn Glu Asp Asp
                85                  90                  95
Ile Asp Ser Ile Pro Pro Leu Glu Ser Asp Glu Glu Asn Asn Thr Glu
            100                 105                 110
Thr Thr Ile Thr Thr Thr Lys Asn Thr Thr Ala Pro Pro Ala Pro Val
        115                 120                 125
Tyr Arg Ser Asn Ser Glu Lys Glu Pro Leu Pro Gln Glu Lys Ser Gln
    130                 135                 140
Lys Gln Pro Asn Gln Val Ser Gly Ser Glu Asn Thr Asp Asn Lys Pro
145                 150                 155                 160
His Ser Glu Gln Ser Val Glu Glu Met
                165
```

<210> 22
<211> 1604
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 22

```
Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1               5                  10                  15
Met Leu Val Ala Ser Cys Leu Gly Asp Ser Arg Pro Gln Lys Ile Trp
            20                  25                  30
Met Ala Pro Ser Leu Thr Glu Ser Asp Met Asp Tyr His Lys Ile Leu
        35                  40                  45
Thr Ala Gly Leu Ser Val Gln Gln Gly Ile Val Arg Gln Arg Val Ile
    50                  55                  60
Pro Val Tyr Gln Val Asn Asn Leu Glu Glu Ile Cys Gln Leu Ile Ile
65                  70                  75                  80
Gln Ala Phe Glu Ala Gly Val Asp Phe Gln Glu Ser Ala Asp Ser Phe
                85                  90                  95
Leu Leu Met Leu Cys Leu His His Ala Tyr Gln Gly Asp Tyr Lys Leu
            100                 105                 110
Phe Leu Glu Ser Gly Ala Val Lys Tyr Leu Glu Gly His Gly Phe Arg
        115                 120                 125
Phe Glu Val Lys Lys Arg Asp Gly Val Lys Arg Leu Glu Glu Leu Leu
    130                 135                 140
Pro Ala Val Ser Ser Gly Lys Asn Ile Lys Arg Thr Leu Ala Ala Met
145                 150                 155                 160
Pro Glu Glu Glu Thr Thr Glu Ala Asn Ala Gly Gln Phe Leu Ser Phe
                165                 170                 175
Ala Ser Leu Phe Leu Pro Lys Leu Val Val Gly Glu Lys Ala Cys Leu
            180                 185                 190
Glu Lys Val Gln Arg Gln Ile Gln Val His Ala Glu Gln Gly Leu Ile
```

```
        195                 200                 205
Gln Tyr Pro Thr Ala Trp Gln Ser Val Gly His Met Met Val Ile Phe
    210                 215                 220
Arg Leu Met Arg Thr Asn Phe Leu Ile Lys Phe Leu Leu Ile His Gln
225                 230                 235                 240
Gly Met His Met Val Ala Gly His Asp Ala Asn Asp Ala Val Ile Ser
                245                 250                 255
Asn Ser Val Ala Gln Ala Arg Phe Ser Gly Leu Leu Ile Val Lys Thr
            260                 265                 270
Val Leu Asp His Ile Leu Gln Lys Thr Glu Arg Gly Val Arg Leu His
        275                 280                 285
Pro Leu Ala Arg Thr Ala Lys Val Lys Asn Glu Val Asn Ser Phe Lys
    290                 295                 300
Ala Ala Leu Ser Ser Leu Ala Lys His Gly Glu Tyr Ala Pro Phe Ala
305                 310                 315                 320
Arg Leu Leu Asn Leu Ser Gly Val Asn Asn Leu Glu His Gly Leu Phe
                325                 330                 335
Pro Gln Leu Ser Ala Ile Ala Leu Gly Val Ala Thr Ala His Gly Ser
            340                 345                 350
Thr Leu Ala Gly Val Asn Val Gly Glu Gln Tyr Gln Gln Leu Arg Glu
        355                 360                 365
Ala Ala Thr Glu Ala Glu Lys Gln Leu Gln Gln Tyr Ala Glu Ser Arg
    370                 375                 380
Glu Leu Asp His Leu Gly Leu Asp Asp Gln Glu Lys Lys Ile Leu Met
385                 390                 395                 400
Asn Phe His Gln Lys Lys Asn Glu Ile Ser Phe Gln Gln Thr Asn Ala
                405                 410                 415
Met Val Thr Leu Arg Lys Glu Arg Leu Ala Lys Leu Thr Glu Ala Ile
            420                 425                 430
Thr Ala Ala Ser Leu Pro Lys Thr Ser Gly His Tyr Asp Asp Asp Asp
        435                 440                 445
Asp Ile Pro Phe Pro Gly Pro Ile Asn Asp Asp Asp Asn Pro Gly His
    450                 455                 460
Gln Asp Asp Asp Pro Thr Asp Ser Gln Asp Thr Thr Ile Pro Asp Val
465                 470                 475                 480
Val Val Asp Pro Asp Asp Gly Ser Tyr Gly Glu Tyr Gln Ser Tyr Ser
                485                 490                 495
Glu Asn Gly Met Asn Ala Pro Asp Asp Leu Val Leu Phe Asp Leu Asp
            500                 505                 510
Glu Asp Asp Glu Asp Thr Lys Pro Val Pro Asn Arg Ser Thr Lys Gly
        515                 520                 525
Gly Gln Gln Lys Asn Ser Gln Lys Gly Gln His Ile Glu Gly Arg Gln
    530                 535                 540
Thr Gln Phe Arg Pro Ile Gln Asn Val Pro Gly Pro His Arg Thr Ile
545                 550                 555                 560
His His Ala Ser Ala Pro Leu Thr Asp Asn Asp Arg Arg Asn Glu Pro
                565                 570                 575
Ser Gly Ser Thr Ser Pro Arg Met Leu Thr Pro Ile Asn Glu Glu Ala
            580                 585                 590
Asp Pro Leu Asp Asp Ala Asp Asp Glu Thr Ser Ser Leu Pro Pro Leu
```

```
        595                 600                 605
Glu Ser Asp Asp Glu Glu Gln Asp Arg Asp Gly Thr Ser Asn Arg Thr
    610                 615                 620
Pro Thr Val Ala Pro Pro Ala Pro Val Tyr Arg Asp His Ser Glu Lys
625                 630                 635                 640
Lys Glu Leu Pro Gln Asp Glu Gln Gln Asp Gln Asp His Thr Gln Glu
                645                 650                 655
Ala Arg Asn Gln Asp Ser Asp Asn Thr Gln Ser Glu His Ser Leu Glu
            660                 665                 670
Glu Met Tyr Arg His Ile Leu Arg Ser Gln Gly Pro Phe Asp Ala Val
        675                 680                 685
Leu Tyr Tyr His Met Met Lys Asp Glu Pro Val Val Phe Ser Thr Ser
    690                 695                 700
Asp Gly Lys Glu Tyr Thr Tyr Pro Asp Ser Leu Glu Glu Glu Tyr Pro
705                 710                 715                 720
Pro Trp Leu Thr Glu Lys Glu Ala Met Asn Glu Glu Asn Arg Phe Val
                725                 730                 735
Thr Leu Asp Gly Gln Gln Phe Tyr Trp Pro Val Met Asn His Lys Asn
            740                 745                 750
Lys Phe Met Ala Ile Leu Gln His His Gln Ala Lys Leu Thr Glu Ala
        755                 760                 765
Ile Thr Thr Ala Ser Lys Ile Lys Val Gly Asp Arg Tyr Pro Asp Asp
    770                 775                 780
Asn Asp Ile Pro Phe Pro Gly Pro Ile Tyr Asp Asp Thr His Pro Asn
785                 790                 795                 800
Pro Ser Asp Asp Asn Pro Asp Asp Ser Arg Asp Thr Thr Ile Pro Gly
                805                 810                 815
Gly Val Val Asp Pro Tyr Asp Asp Glu Ser Asn Asn Tyr Pro Asp Tyr
            820                 825                 830
Glu Asp Ser Ala Glu Gly Thr Thr Gly Asp Leu Asp Leu Phe Asn Leu
        835                 840                 845
Asp Asp Asp Asp Asp Asp Ser Arg Pro Gly Pro Pro Asp Arg Gly Gln
    850                 855                 860
Asn Lys Glu Arg Ala Ala Arg Thr Tyr Gly Leu Gln Asp Pro Thr Leu
865                 870                 875                 880
Asp Gly Ala Lys Lys Val Pro Glu Leu Thr Pro Gly Ser His Gln Pro
                885                 890                 895
Gly Asn Leu His Ile Thr Lys Ser Gly Ser Asn Thr Asn Gln Pro Gln
            900                 905                 910
Gly Asn Met Ser Ser Thr Leu His Ser Met Thr Pro Ile Gln Glu Glu
        915                 920                 925
Ser Glu Pro Asp Asp Gln Lys Asp Asn Asp Asp Glu Ser Leu Thr Ser
    930                 935                 940
Leu Asp Ser Glu Gly Asp Glu Asp Gly Glu Ser Ile Ser Glu Glu Asn
945                 950                 955                 960
Thr Pro Thr Val Ala Pro Pro Ala Pro Val Tyr Lys Asp Thr Gly Val
                965                 970                 975
Asp Thr Asn Gln Gln Asn Gly Pro Ser Ser Thr Val Asp Ser Gln Gly
            980                 985                 990
Ser Glu Ser Glu Ala Leu Pro Ile Asn Ser Lys Lys Ser Ser Ala Leu
```

```
        995                 1000                1005
Glu Glu Thr Tyr Tyr His Leu Leu Lys Thr Gln Gly Pro Phe Glu Ala
    1010                1015                1020
Ile Asn Tyr Tyr His Leu Met Ser Asp Glu Pro Ile Ala Phe Ser Thr
1025                1030                1035                1040
Glu Ser Gly Lys Glu Tyr Ile Phe Pro Asp Ser Leu Glu Glu Ala Tyr
                1045                1050                1055
Pro Pro Trp Leu Ser Glu Lys Glu Ala Leu Glu Lys Glu Asn Arg Tyr
            1060                1065                1070
Leu Val Ile Asp Gly Gln Gln Phe Leu Trp Pro Val Met Ser Leu Arg
        1075                1080                1085
Asp Lys Phe Leu Ala Val Leu Gln His Asp Ala Lys Leu Thr Glu Ala
    1090                1095                1100
Ile Thr Ser Thr Ser Ile Leu Lys Thr Gly Arg Arg Tyr Asp Asp Asp
1105                1110                1115                1120
Asn Asp Ile Pro Phe Pro Gly Pro Ile Asn Asp Asn Glu Asn Ser Gly
                1125                1130                1135
Gln Asn Asp Asp Asp Pro Thr Asp Ser Gln Asp Thr Thr Ile Pro Asp
            1140                1145                1150
Val Ile Ile Asp Pro Asn Asp Gly Gly Tyr Asn Asn Tyr Ser Asp Tyr
        1155                1160                1165
Ala Asn Asp Ala Ala Ser Ala Pro Asp Asp Leu Val Leu Phe Asp Leu
    1170                1175                1180
Glu Asp Glu Asp Asp Ala Asp Asn Pro Ala Gln Asn Thr Pro Glu Lys
1185                1190                1195                1200
Asn Asp Arg Pro Ala Thr Thr Lys Leu Arg Asn Gly Gln Asp Gln Asp
                1205                1210                1215
Gly Asn Gln Gly Glu Thr Ala Ser Pro Arg Val Ala Pro Asn Gln Tyr
            1220                1225                1230
Arg Asp Lys Pro Met Pro Gln Val Gln Asp Arg Ser Glu Asn His Asp
        1235                1240                1245
Gln Thr Leu Gln Thr Gln Ser Arg Val Leu Thr Pro Ile Ser Glu Glu
    1250                1255                1260
Ala Asp Pro Ser Asp His Asn Asp Gly Asp Asn Glu Ser Ile Pro Pro
1265                1270                1275                1280
Leu Glu Ser Asp Asp Glu Gly Ser Thr Asp Thr Thr Ala Ala Glu Thr
                1285                1290                1295
Lys Pro Ala Thr Ala Pro Pro Ala Pro Val Tyr Arg Ser Ile Ser Val
            1300                1305                1310
Asp Asp Ser Val Pro Ser Glu Asn Ile Pro Ala Gln Ser Asn Gln Thr
        1315                1320                1325
Asn Asn Glu Asp Asn Val Arg Asn Asn Ala Gln Ser Glu Gln Ser Ile
    1330                1335                1340
Ala Glu Met Tyr Gln His Ile Leu Lys Thr Gln Gly Pro Phe Asp Ala
1345                1350                1355                1360
Ile Leu Tyr Tyr His Met Met Lys Glu Glu Pro Ile Ile Phe Ser Thr
                1365                1370                1375
Ser Asp Gly Lys Glu Tyr Thr Tyr Pro Asp Ser Leu Glu Asp Glu Tyr
            1380                1385                1390
Pro Pro Trp Leu Ser Glu Lys Glu Ala Met Asn Glu Asp Asn Arg Phe
```

```
        1395                1400                1405
Ile Thr Met Asp Gly Gln Gln Phe Tyr Trp Pro Val Met Asn His Arg
    1410                1415                1420
Asn Lys Phe Met Ala Ile Leu Gln His His Arg Leu Val Leu Phe Asp
1425                1430                1435                1440
Leu Glu Asp Gly Asp Glu Asp Asp His Arg Pro Ser Ser Ser Ser Glu
                1445                1450                1455
Asn Asn Asn Lys His Ser Leu Thr Gly Thr Asp Ser Asn Lys Thr Ser
            1460                1465                1470
Asn Trp Asn Arg Asn Pro Thr Asn Met Pro Lys Lys Asp Ser Thr Gln
        1475                1480                1485
Asn Asn Asp Asn Pro Ala Gln Arg Ala Gln Glu Tyr Ala Arg Asp Asn
    1490                1495                1500
Ile Gln Asp Thr Pro Thr Pro His Arg Ala Leu Thr Pro Ile Ser Glu
1505                1510                1515                1520
Glu Thr Gly Ser Asn Gly His Asn Glu Asp Asp Ile Asp Ser Ile Pro
                1525                1530                1535
Pro Leu Glu Ser Asp Glu Glu Asn Asn Thr Glu Thr Thr Ile Thr Thr
            1540                1545                1550
Thr Lys Asn Thr Thr Ala Pro Pro Ala Pro Val Tyr Arg Ser Asn Ser
        1555                1560                1565
Glu Lys Glu Pro Leu Pro Gln Glu Lys Ser Gln Lys Gln Pro Asn Gln
    1570                1575                1580
Val Ser Gly Ser Glu Asn Thr Asp Asn Lys Pro His Ser Glu Gln Ser
1585                1590                1595                1600
Val Glu Glu Met
```

<210> 23
<211> 4815
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 23
atgcccatgg ggtctctgca accgctggcc accttgtacc tgctggggat gctggtcgct 60
tcctgcctcg gagattctcg tcctcagaaa atctggatgg cgccgagtct cactgaatct 120
gacatggatt accacaagat cttgacagca ggtctgtccg ttcaacaggg gattgttcgg 180
caaagagtca tcccagtgta tcaagtaaac aatcttgaag aaatttgcca acttatcata 240
caggcctttg aagcaggtgt tgattttcaa gagagtgcgg acagtttcct tctcatgctt 300
tgtcttcatc atgcgtacca gggagattac aaacttttct tggaaagtgg cgcagtcaag 360
tatttggaag ggcacgggtt ccgttttgaa gtcaagaagc gtgatggagt gaagcgcctt 420
gaggaattgc tgccagcagt atctagtgga aaaaacatta agagaacact tgctgccatg 480
ccggaagagg agacaactga agctaatgcc ggtcagtttc tctcctttgc aagtctattc 540
cttccgaaat tggtagtagg agaaaaggct tgccttgaga aggttcaaag gcaaattcaa 600
gtacatgcag agcaaggact gatacaatat ccaacagctt ggcaatcagt aggacacatg 660
atggtgattt tccgtttgat gcgaacaaat tttctgatca aatttctcct aatacaccaa 720

gggatgcaca tggttgccgg gcatgatgcc aacgatgctg tgatttcaaa ttcagtggct 780
caagctcgtt tttcaggctt attgattgtc aaaacagtac ttgatcatat cctacaaaag 840
acagaacgag gagttcgtct ccatcctctt gcaaggaccg ccaaggtaaa aaatgaggtg 900
aactccttta aggctgcact cagctccctg gccaagcatg gagagtatgc tcctttcgcc 960
cgacttttga acctttctgg agtaaataat cttgagcatg gtcttttccc tcaactatcg 1020
gcaattgcac tcggagtcgc cacagcacac gggagtaccc tcgcaggagt aaatgttgga 1080
gaacagtatc aacaactcag agaggctgcc actgaggctg agaagcaact ccaacaatat 1140
gcagagtctc gcgaacttga ccatcttgga cttgatgatc aggaaaagaa aattcttatg 1200
aacttccatc agaaaaagaa cgaaatcagc ttccagcaaa caaacgctat ggtaactcta 1260
agaaaagagc gcctggccaa gctgacagaa gctatcactg ctgcgtcact gcccaaaaca 1320
agtggacatt acgatgatga tgacgacatt ccctttccag gacccatcaa tgatgacgac 1380
aatcctggcc atcaagatga tgatccgact gactcacagg atacgaccat tcccgatgtg 1440
gtggttgatc ctgatgatgg aagctacggc gaataccaga gttactcgga aaacggcatg 1500
aatgcaccag atgacttggt cctattcgat ctagacgagg acgacgagga cactaagcca 1560
gtgcctaata gatcgaccaa gggtggacaa cagaagaaca gtcaaaaggg ccagcatata 1620
gagggcagac agacacaatt caggccaatt caaaatgtcc caggccctca cagaacaatc 1680
caccacgcca gtgcgccact cacggacaat gacagaagaa atgaaccctc cggctcaacc 1740
agccctcgca tgctgacacc aattaacgaa gaggcagacc cactggacga tgccgacgac 1800
gagacgtcta gccttccgcc cttggagtca gatgatgaag agcaggacag ggacggaact 1860
tccaaccgca cacccactgt cgccccaccg gctcccgtat acagagatca ctctgaaaag 1920
aaagaactcc cgcaagacga gcaacaagat caggaccaca ctcaagaggc caggaaccag 1980
gacagtgaca acacccagtc agaacactcc cttgaggaga tgtatcgcca cattctaaga 2040
tcacaggggc catttgatgc tgttttgtat tatcatatga tgaaggatga gcctgtagtt 2100
ttcagtacca gtgatggcaa agagtacacg tatccagact cccttgaaga ggaatatcca 2160
ccatggctca ctgaaaaaga ggctatgaat gaagagaata gatttgttac attggatggt 2220
caacaatttt attggccggt gatgaatcac aagaataaat tcatggcaat cctgcaacat 2280
catcaggcta aattgaccga agccatcacg actgcatcga agatcaaggt tggagaccgt 2340
tatcctgatg acaatgatat tccatttccc gggccgatct atgatgacac tcaccccaat 2400
ccctctgatg acaatcctga tgattcacgt gatacaacta ttccaggtgg tgttgttgac 2460
ccgtatgatg atgagagtaa taattatcct gactacgagg attcggctga aggcaccaca 2520
ggagatcttg atctcttcaa tttggacgac gacgatgatg acagccgacc aggaccacca 2580
gacagggggc agaacaagga gagggcggcc cggacatatg gcctccaaga tccgaccttg 2640
gacggagcga aaaaggtgcc ggagttgacc ccaggttccc atcaaccagg caacctccac 2700
atcaccaagt cgggttcaaa caccaaccaa ccacaaggca atatgtcatc tactctccat 2760
agtatgaccc ctatacagga agaatcagag cccgatgatc aaaaagataa tgatgacgag 2820
agtctcacat cccttgactc tgaaggtgac gaagatggtg agagcatctc tgaggagaac 2880
accccaactg tagctccacc agcaccagtc tacaaagaca ctggagtaga cactaatcag 2940
cagaatggac caagcagtac tgtagatagt caaggttctg aaagtgaagc tctcccaatc 3000
aactctaaaa agagttccgc actagaagaa acatattatc atctcctaaa aacacagggt 3060
ccatttgagg caatcaatta ttatcaccta atgagtgatg aacccattgc ttttagcact 3120
gaaagtggca aggaatatat ctttccagac tcccttgaag aagcctaccc gccgtggttg 3180
agtgagaagg aggccttaga gaaggaaaat cgttatctgg tcattgatgg ccagcaattc 3240
ctctggccgg taatgagcct acgggacaag ttccttgccg ttcttcaaca tgacgccaaa 3300
ttgaccgaag ctattacttc cacctctatc ctcaaaacag gaaggcggta tgatgatgac 3360
aatgacatac cctttccagg gccaatcaat gataacgaga actctggtca gaacgatgac 3420
gatccaacag actcccagga taccacaatc ccggatgtaa taatcgatcc aaacgatggt 3480
gggtataata attacagcga ttatgcaaat gatgctgcaa gtgctcctga tgacctagtt 3540
ctttttgacc ttgaggacga ggatgatgct gataacccgg ctcaaaacac gccagaaaaa 3600
aatgatagac cagcaacaac aaagctgaga aatggacagg accaggatgg aaaccaaggc 3660
gaaactgcat ccccacgggt agcccccaac caatacagag acaagccaat gccacaagta 3720

caggacagat ccgaaaatca tgaccaaacc cttcaaacac agtccagggt tttgactcct 3780
atcagcgagg aagcagaccc cagcgaccac aacgatggtg acaatgaaag cattcctccc 3840
ctggaatcag acgacgaggg tagcactgat actactgcag cagaaacaaa gcctgccact 3900
gcacctcccg ctcccgtcta ccgaagtatc tccgtagatg attctgtccc ctcagagaac 3960
attcccgcac agtccaatca aacgaacaat gaggacaatg tcaggaacaa tgctcagtcg 4020
gagcaatcca ttgcagaaat gtatcaacat atcttgaaaa cacaaggacc ttttgatgcc 4080
atcctttact accatatgat gaaagaagag cccatcattt tcagcactag tgatgggaag 4140
gagtatacat atccagactc tcttgaagat gagtatccac cctggctcag cgagaaggaa 4200
gccatgaacg aagacaatag attcataacc atggatggtc agcagtttta ctggcctgtg 4260
atgaatcata gaaataaatt catggcaatc ctccagcatc acaggcttgt tcttttgac 4320
cttgaagatg gtgacgagga tgatcaccga ccgtcaagtt catcagagaa caacaacaaa 4380
cacagtctta caggaactga cagtaacaaa acaagtaact ggaatcgaaa cccgactaat 4440
atgccaaaga aagactccac acaaaacaat gacaatcctg cacagcgggc tcaagaatac 4500
gccagggata acatccagga tacaccaaca ccccatcgag ctctaactcc catcagcgaa 4560
gaaaccggct ccaatggtca caatgaagat gacattgata gcatccctcc tttggaatca 4620
gacgaagaaa acaacactga gacaaccatt accaccacaa aaaataccac tgctccacca 4680
gcacctgttt atcggagtaa ttcagaaaag gagcccctcc cgcaagaaaa atcccagaag 4740
caaccaaacc aagtgagtgg tagtgagaat accgacaata aacctcactc agagcaatca 4800
gtggaagaaa tgtaa 4815

<210> 24
<211> 744
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 24
atgcccatgg ggtctctgca accgctggcc accttgtacc tgctggggat gctggtcgct 60
tcctgcctcg gaagagtcca accaacagaa tctattgtta gatttcctaa tattacaaac 120
ttgtgccctt ttgatgaagt ttttaacgcc accagatttg catctgttta tgcttggaac 180
aggaagagaa tcagcaactg tgttgctgat tattctgtcc tatataatct cgcaccattt 240
ttcactttta agtgttatgg agtgtctcct actaaattaa atgatctctg ctttactaat 300
gtctatgcag attcatttgt aattagaggt gatgaagtca gacaaatcgc tccagggcaa 360
actggaaaca ttgctgatta taattataaa ttaccagatg attttacagg ctgcgttata 420
gcttggaatt ctaacaagct tgattctaag gttagtggta attataatta cctgtataga 480
ttgtttagga agtctaatct caaacctttt gagagagata tttcaactga aatctatcag 540
gccggtaaca aaccttgtaa tggtgttgca ggttttaatt gttactttcc tttacgatca 600
tatagtttcc gacccactta tggtgttggt caccaaccat acagagtagt agtactttct 660
tttgaacttc tacatgcacc agcaactgtt tgtggaccta aaaagtctac taatttggtt 720
aaaaacaaat gtgtcaattt ctaa 744

<210> 25
<211> 247
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 25
Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1   5   10   15
Met Leu Val Ala Ser Cys Leu Gly Arg Val Gln Pro Thr Glu Ser Ile
            20   25   30
Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Asp Glu Val Phe
        35   40   45
Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile
    50   55   60
Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Leu Ala Pro Phe
65   70   75   80
Phe Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu
                85   90   95
Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu
            100   105   110
Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Asn Ile Ala Asp Tyr Asn
        115   120   125
Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser
    130   135   140
Asn Lys Leu Asp Ser Lys Val Ser Gly Asn Tyr Asn Tyr Leu Tyr Arg
145   150   155   160
Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr
                165   170   175
Glu Ile Tyr Gln Ala Gly Asn Lys Pro Cys Asn Gly Val Ala Gly Phe
            180   185   190
Asn Cys Tyr Phe Pro Leu Arg Ser Tyr Ser Phe Arg Pro Thr Tyr Gly
        195   200   205
Val Gly His Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu
    210   215   220
His Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val
225   230   235   240
Lys Asn Lys Cys Val Asn Phe
                245

<210> 26
<211> 4056
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 26
atgcccatgg ggtctctgca accgctggcc accttgtacc tgctggggat gctggtcgct 60
tcctgcctcg gaagagtcca accaacagaa tctattgtta gatttcctaa tattacaaac 120
ttgtgccctt ttggtgaagt ttttaacgcc accagatttg catctgttta tgcttggaac 180
aggaagagaa tcagcaactg tgttgctgat tattctgtcc tatataattc cgcatcattt 240
tccactttta agtgttatgg agtgtctcct actaaattaa atgatctctg ctttactaat 300

gtctatgcag attcatttgt aattagaggt gatgaagtca gacaaatcgc tccagggcaa 360
actggaaaga ttgctgatta taattataaa ttaccagatg attttacagg ctgcgttata 420
gcttggaatt ctaacaatct tgattctaag gttggtggta attataatta ccggtataga 480
ttgtttagga agtctaatct caaacctttt gagagagata tttcaactga aatctatcag 540
gccggtagca aaccttgtaa tggtgttgaa ggttttaatt gttactttcc tttacaatca 600
tatggtttcc aacccactaa tggtgttggt taccaaccat acagagtagt agtactttct 660
tttgaacttc tacatgcacc agcaactgtt tgtggaccta aaaagtctac taatttggtt 720
aaaaacaaat gtgtcaattt ctcaggtggt ggcggttcag gcggaggtgg ctctggcggt 780
ggcggatcga tgtctgataa tggaccccaa aatcagcgaa atgcaccccg cattacgttt 840
ggtggaccct cagattcaac tggcagtaac cagaatggag aacgcagtgg ggcgcgatca 900
aaacaacgtc ggccccaagg tttacccaat aatactgcgt cttggttcac cgctctcact 960
caacatggca aggaaggcct taaattccct cgaggacaag gcgttccaat taacaccaat 1020
agcagtccag atgaccaaat tggctactac cgaagagcta ccagacgaat tcgtggtggt 1080
gacggtaaaa tgaaagatct cagtccaaga tggtatttct actacctagg aactgggcca 1140
gaagctggac ttccctatgg tgctaacaaa gacggcatca tatgggttgc aactgaggga 1200
gccttgaata caccaaaaga tcacattggc acccgcaatc ctgctaacaa tgctgcaatc 1260
gtgctacaac ttcctcaagg aacaacattg ccaaaaggct tctacgcaga agggagcaga 1320
ggcggcagtc aagcctcttc tcgttcctca tcacgtagtc gcaacagttc aagaaattca 1380
actccaggca gcagtatggg aacttctcct gctagaatgg ctggcaatgg ctgtgatgct 1440
gctcttgctt tgctgctgct tgacagattg aaccagcttg agagcaaaat gtctggtaaa 1500
ggccaacaac aacaaggcca aactgtcact aagaaatctg ctgctgaggc ttctaagaag 1560
cctcggcaaa aacgtactgc cactaaagca tacaatgtaa cacaagcttt cggcagacgt 1620
ggtccagaac aaacccaagg aaattttggg gaccaggaac taatcagaca aggaactgat 1680
tacaaacatt ggccgcaaat tgcacaattt gcccccagcg cttcagcgtt cttcggaatg 1740
tcgcgcattg gcatggaagt cacaccttcg ggaacgtggt tgacctacac aggtgccatc 1800
aaattggatg acaaagatcc aaatttcaaa gatcaagtca ttttgctgaa taagcatatt 1860
gacgcataca aaacattccc accaacagag cctaaaaagg acaaaaagaa gaaggcttat 1920
gaaactcaag ccttaccgca gagacagaag aaacagcaaa ctgtgactct tcttcctgct 1980
gcagatttgg atgatttctc caaacaattg caacaatcca tgagcagtgc tgactcaact 2040
caggccatgg catcccctgc tgcacctcgt gctgtttcct ttgccgataa caatgatata 2100
acaaatacaa acctatctcg aggtagagga cgtaatccaa aaccacgagc tgcaccaaat 2160
aacactgtct cttggtacac tgggcttacc caacacggga aagtccctct tacctttcca 2220
cctgggcagg gtgtacctct taatgccaat tctacccctg cgcaaaatgc tgggtattgg 2280
cggagacagg acagaaaaat taataccggg aatggaatta agcaactggc tcccaggtgg 2340
tacttctact acactggaac tggacccgaa gcagcactcc cattccgggc tgttaaggat 2400
ggcatcgttt gggtccatga agatggcgcc actgatgctc cttcaacttt tgggacgcgg 2460
aaccctaaca atgattcagc tattgttaca caattcgcgc ccggtactaa gcttcctaaa 2520
aacttccaca ttgaggggac tggaggcaat agtcaatcat cttcaagagc ctctagctta 2580
agcagaaact cttccagatc tagttcacaa ggttcaagat caggaaactc tacccgcggc 2640
acttctccag gtccatctgg aatcggagca gtaggaggtg atctacttta ccttgatctt 2700
ctgaacagac tacaagccct tgagtctggc aaagtaaagc aatcgcagcc aaaagtaatc 2760
actaagaaag atgctgctgc tgctaaaaat aagatgcgcc acaagcgcac ttccaccaaa 2820
agtttcaaca tggtgcaagc ttttggtctt cgcggaccag gagacctcca gggaaacttt 2880
ggtgatcttc aattgaataa actcggcact gaggacccac gttggcccca aattgctgag 2940
cttgctccta cagccagtgc ttttatgggt atgtcgcaat ttaaacttac ccatcagaac 3000
aatgatgatc atggcaaccc tgtgtacttc cttcggtaca gtggagccat taaacttgac 3060
ccaaagaatc ccaactacaa taagtggttg gagcttcttg agcaaaatat tgatgcctac 3120
aaaaccttcc ctaagaagga aaagaaacaa aaggcaccaa aagaagaatc aacagaccaa 3180
atgtctgaac ctccaaagga gcagcgtgtg caaggtagca tcactcagcg cactcgcacc 3240
cgtccaagtg ttcagcctgg tccaatgatt gatgttaaca ctgattctgg tggcggtggc 3300

tcgggcggag gtgggtcggg tggcggcgga tcagaagcaa aaccttctgg ctcagttgtg 3360
gaacaggctg aaggtgttga atgtgatttt tcacctcttc tgtctggcac acctcctcag 3420
gtttataatt tcaagcgttt ggtttttacc aattgcaatt ataatcttac caaattgctt 3480
tcacttttt ctgtgaatga ttttacttgt agtcaaatat ctccagcagc aattgctagc 3540
aactgttatt cttcactgat tttggattac ttttcatacc cacttagtat gaaatccgat 3600
ctcagtgtta gttctgctgg tccaatatcc cagtttaatt ataaacagtc cttttctaat 3660
cccacatgtt tgattttagc gactgttcct cataacctta ctactattac taagcctctt 3720
aagtacagct atattaacaa gtgctctcgt cttctttctg atgatcgtac tgaagtacct 3780
cagttagtga acgctaatca atactcaccc tgtgtatcca ttgtcccatc cactgtgtgg 3840
gaagacggtg attattatag gaaacaacta tctccacttg aaggtggtgg ctggcttgtt 3900
gctagtggct caactgttgc catgactgag caattacaga tgggctttgg tattacagtt 3960
caatatggta cagacaccaa tagtgtttgc cccaagcttg aatttgctaa tgacacaaaa 4020
attgcctctc aattaggcaa ttgcgtggaa tattag 4056

<210> 27
<211> 1351
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 27
Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1   5   10   15
Met Leu Val Ala Ser Cys Leu Gly Arg Val Gln Pro Thr Glu Ser Ile
            20   25   30
Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe
        35   40   45
Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile
    50   55   60
Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe
65   70   75   80
Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu
                85   90   95
Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu
            100   105   110
Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn
        115   120   125
Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser
    130   135   140
Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Arg Tyr Arg
145   150   155   160
Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr
                165   170   175
Glu Ile Tyr Gln Ala Gly Ser Lys Pro Cys Asn Gly Val Glu Gly Phe
            180   185   190
Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly
        195   200   205
Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu

```
    210                 215                 220
His Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val
225                 230                 235                 240
Lys Asn Lys Cys Val Asn Phe Ser Gly Gly Gly Gly Ser Gly Gly Gly
                245                 250                 255
Gly Ser Gly Gly Gly Gly Ser Met Ser Asp Asn Gly Pro Gln Asn Gln
            260                 265                 270
Arg Asn Ala Pro Arg Ile Thr Phe Gly Gly Pro Ser Asp Ser Thr Gly
        275                 280                 285
Ser Asn Gln Asn Gly Glu Arg Ser Gly Ala Arg Ser Lys Gln Arg Arg
    290                 295                 300
Pro Gln Gly Leu Pro Asn Asn Thr Ala Ser Trp Phe Thr Ala Leu Thr
305                 310                 315                 320
Gln His Gly Lys Glu Gly Leu Lys Phe Pro Arg Gly Gln Gly Val Pro
                325                 330                 335
Ile Asn Thr Asn Ser Ser Pro Asp Asp Gln Ile Gly Tyr Tyr Arg Arg
            340                 345                 350
Ala Thr Arg Arg Ile Arg Gly Gly Asp Gly Lys Met Lys Asp Leu Ser
        355                 360                 365
Pro Arg Trp Tyr Phe Tyr Tyr Leu Gly Thr Gly Pro Glu Ala Gly Leu
    370                 375                 380
Pro Tyr Gly Ala Asn Lys Asp Gly Ile Ile Trp Val Ala Thr Glu Gly
385                 390                 395                 400
Ala Leu Asn Thr Pro Lys Asp His Ile Gly Thr Arg Asn Pro Ala Asn
                405                 410                 415
Asn Ala Ala Ile Val Leu Gln Leu Pro Gln Gly Thr Thr Leu Pro Lys
            420                 425                 430
Gly Phe Tyr Ala Glu Gly Ser Arg Gly Gly Ser Gln Ala Ser Ser Arg
        435                 440                 445
Ser Ser Ser Arg Ser Arg Asn Ser Ser Arg Asn Ser Thr Pro Gly Ser
    450                 455                 460
Ser Met Gly Thr Ser Pro Ala Arg Met Ala Gly Asn Gly Cys Asp Ala
465                 470                 475                 480
Ala Leu Ala Leu Leu Leu Leu Asp Arg Leu Asn Gln Leu Glu Ser Lys
                485                 490                 495
Met Ser Gly Lys Gly Gln Gln Gln Gln Gly Gln Thr Val Thr Lys Lys
            500                 505                 510
Ser Ala Ala Glu Ala Ser Lys Lys Pro Arg Gln Lys Arg Thr Ala Thr
        515                 520                 525
Lys Ala Tyr Asn Val Thr Gln Ala Phe Gly Arg Arg Gly Pro Glu Gln
    530                 535                 540
Thr Gln Gly Asn Phe Gly Asp Gln Glu Leu Ile Arg Gln Gly Thr Asp
545                 550                 555                 560
Tyr Lys His Trp Pro Gln Ile Ala Gln Phe Ala Pro Ser Ala Ser Ala
                565                 570                 575
Phe Phe Gly Met Ser Arg Ile Gly Met Glu Val Thr Pro Ser Gly Thr
            580                 585                 590
Trp Leu Thr Tyr Thr Gly Ala Ile Lys Leu Asp Asp Lys Asp Pro Asn
        595                 600                 605
Phe Lys Asp Gln Val Ile Leu Leu Asn Lys His Ile Asp Ala Tyr Lys
```

```
    610                 615                 620
Thr Phe Pro Pro Thr Glu Pro Lys Lys Asp Lys Lys Lys Lys Ala Tyr
625                 630                 635                 640
Glu Thr Gln Ala Leu Pro Gln Arg Gln Lys Lys Gln Gln Thr Val Thr
                645                 650                 655
Leu Leu Pro Ala Ala Asp Leu Asp Asp Phe Ser Lys Gln Leu Gln Gln
            660                 665                 670
Ser Met Ser Ser Ala Asp Ser Thr Gln Ala Met Ala Ser Pro Ala Ala
        675                 680                 685
Pro Arg Ala Val Ser Phe Ala Asp Asn Asn Asp Ile Thr Asn Thr Asn
    690                 695                 700
Leu Ser Arg Gly Arg Gly Arg Asn Pro Lys Pro Arg Ala Ala Pro Asn
705                 710                 715                 720
Asn Thr Val Ser Trp Tyr Thr Gly Leu Thr Gln His Gly Lys Val Pro
                725                 730                 735
Leu Thr Phe Pro Pro Gly Gln Gly Val Pro Leu Asn Ala Asn Ser Thr
            740                 745                 750
Pro Ala Gln Asn Ala Gly Tyr Trp Arg Arg Gln Asp Arg Lys Ile Asn
        755                 760                 765
Thr Gly Asn Gly Ile Lys Gln Leu Ala Pro Arg Trp Tyr Phe Tyr Tyr
    770                 775                 780
Thr Gly Thr Gly Pro Glu Ala Ala Leu Pro Phe Arg Ala Val Lys Asp
785                 790                 795                 800
Gly Ile Val Trp Val His Glu Asp Gly Ala Thr Asp Ala Pro Ser Thr
                805                 810                 815
Phe Gly Thr Arg Asn Pro Asn Asn Asp Ser Ala Ile Val Thr Gln Phe
            820                 825                 830
Ala Pro Gly Thr Lys Leu Pro Lys Asn Phe His Ile Glu Gly Thr Gly
        835                 840                 845
Gly Asn Ser Gln Ser Ser Ser Arg Ala Ser Ser Leu Ser Arg Asn Ser
    850                 855                 860
Ser Arg Ser Ser Ser Gln Gly Ser Arg Ser Gly Asn Ser Thr Arg Gly
865                 870                 875                 880
Thr Ser Pro Gly Pro Ser Gly Ile Gly Ala Val Gly Gly Asp Leu Leu
                885                 890                 895
Tyr Leu Asp Leu Leu Asn Arg Leu Gln Ala Leu Glu Ser Gly Lys Val
            900                 905                 910
Lys Gln Ser Gln Pro Lys Val Ile Thr Lys Lys Asp Ala Ala Ala Ala
        915                 920                 925
Lys Asn Lys Met Arg His Lys Arg Thr Ser Thr Lys Ser Phe Asn Met
    930                 935                 940
Val Gln Ala Phe Gly Leu Arg Gly Pro Gly Asp Leu Gln Gly Asn Phe
945                 950                 955                 960
Gly Asp Leu Gln Leu Asn Lys Leu Gly Thr Glu Asp Pro Arg Trp Pro
                965                 970                 975
Gln Ile Ala Glu Leu Ala Pro Thr Ala Ser Ala Phe Met Gly Met Ser
            980                 985                 990
Gln Phe Lys Leu Thr His Gln Asn Asn Asp Asp His Gly Asn Pro Val
        995                 1000                1005
Tyr Phe Leu Arg Tyr Ser Gly Ala Ile Lys Leu Asp Pro Lys Asn Pro
```

```
    1010                1015                1020
Asn Tyr Asn Lys Trp Leu Glu Leu Leu Glu Gln Asn Ile Asp Ala Tyr
1025                1030                1035                1040
Lys Thr Phe Pro Lys Lys Glu Lys Lys Gln Lys Ala Pro Lys Glu Glu
                1045                1050                1055
Ser Thr Asp Gln Met Ser Glu Pro Pro Lys Glu Gln Arg Val Gln Gly
            1060                1065                1070
Ser Ile Thr Gln Arg Thr Arg Thr Arg Pro Ser Val Gln Pro Gly Pro
        1075                1080                1085
Met Ile Asp Val Asn Thr Asp Ser Gly Gly Gly Gly Ser Gly Gly Gly
    1090                1095                1100
Gly Ser Gly Gly Gly Gly Ser Glu Ala Lys Pro Ser Gly Ser Val Val
1105                1110                1115                1120
Glu Gln Ala Glu Gly Val Glu Cys Asp Phe Ser Pro Leu Leu Ser Gly
                1125                1130                1135
Thr Pro Pro Gln Val Tyr Asn Phe Lys Arg Leu Val Phe Thr Asn Cys
            1140                1145                1150
Asn Tyr Asn Leu Thr Lys Leu Leu Ser Leu Phe Ser Val Asn Asp Phe
        1155                1160                1165
Thr Cys Ser Gln Ile Ser Pro Ala Ala Ile Ala Ser Asn Cys Tyr Ser
    1170                1175                1180
Ser Leu Ile Leu Asp Tyr Phe Ser Tyr Pro Leu Ser Met Lys Ser Asp
1185                1190                1195                1200
Leu Ser Val Ser Ser Ala Gly Pro Ile Ser Gln Phe Asn Tyr Lys Gln
                1205                1210                1215
Ser Phe Ser Asn Pro Thr Cys Leu Ile Leu Ala Thr Val Pro His Asn
            1220                1225                1230
Leu Thr Thr Ile Thr Lys Pro Leu Lys Tyr Ser Tyr Ile Asn Lys Cys
        1235                1240                1245
Ser Arg Leu Leu Ser Asp Asp Arg Thr Glu Val Pro Gln Leu Val Asn
    1250                1255                1260
Ala Asn Gln Tyr Ser Pro Cys Val Ser Ile Val Pro Ser Thr Val Trp
1265                1270                1275                1280
Glu Asp Gly Asp Tyr Tyr Arg Lys Gln Leu Ser Pro Leu Glu Gly Gly
                1285                1290                1295
Gly Trp Leu Val Ala Ser Gly Ser Thr Val Ala Met Thr Glu Gln Leu
            1300                1305                1310
Gln Met Gly Phe Gly Ile Thr Val Gln Tyr Gly Thr Asp Thr Asn Ser
        1315                1320                1325
Val Cys Pro Lys Leu Glu Phe Ala Asn Asp Thr Lys Ile Ala Ser Gln
    1330                1335                1340
Leu Gly Asn Cys Val Glu Tyr
1345                1350


<210> 28
<211> 4098
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Synthetic Construct

<400> 28
atgcctatgg gcagcctgca gccactggct acactgtacc tgctgggcat gctggtggcc 60
tcttgtctgg gcataggaga gtgtccaaag tatgtcagga gtgcaaaatt aaggatggtt 120
acaggactaa ggaacatccc atccattcaa tccagaggtt tgtttggagc cattgccggt 180
ttcattgaag gggggtggac tggaatggta gatgggtggt atggttatca tcatcagaat 240
gagcaaggat ctggctatgc tgcagatcaa aaaagtacac aaaatgccat taacgggatt 300
acaaacaagg tgaattctgt aattgagaaa atgaacactc aattcacagc tgtgggcaaa 360
gaattcaaca aattggaaag aaggatggaa aacttaaata aaaaagttga tgatgggttt 420
ctagacattt ggacatataa tgcagaattg ttggttctac tggaaaatga aaggactttg 480
gatttccatg actccgccag ccaaggcact aagagaagct acgagcagat ggaaaccgga 540
ggcgaacggc agaacgccac agagatcaga gcctctgtgg gccgtatggt cggcggcatc 600
ggcagattct acatccagat gtgcaccgaa ctgaagctga gcgactacga gggccgcctg 660
atccagaaca gcatcacaat cgagagaatg gtgctgtccg cctttgacga gcggagaaac 720
aaatacctgg aagagcaccc tagcgccgga aaagatccta agaaaaccgg cggacctatc 780
tacagaagaa gagatggtaa gtgggtgaga gagctgattc tgtacgataa ggaagagatt 840
cgaagaatct ggagacaggc caacaacggc gaggatgcca ccgcaggcct gacacacctg 900
atgatctggc acagcaacct gaacgatgcg acctaccagc gcacgcgggc cctggtcaga 960
accggcatgg atcctcggat gtgtagcctg atgcagggca gcacactgcc aagacggagt 1020
ggggccgccg gcgctgcagt gaagggcgtc ggaaccatgg tgatggagct gatccggatg 1080
ataaagcggg gcatcaacga cagaaacttc tggcgaggcg agaacggccg aagaacccgg 1140
atcgcctacg agagaatgtg caacatcctg aaaggaaaat tccagaccgc cgcccagcgg 1200
gccatgatgg accaggtgcg cgagagcaga aaccccggca atgccgagat cgaggacctg 1260
atcttcctgg ccagaagcgc cctcattctt agaggctctg tggcccacaa gagctgtctg 1320
cctgcctgtg tgtacggcct ggcagtggcc tcaggctacg acttcgagcg ggaaggatac 1380
agtctggtgg gcatcgaccc tttcagactc ctgcagaata gccaggtgtt tagcctgatc 1440
agaccaaacg aaaaccccgc ccataagagc cagctggtgt ggatggcctg ccacagcgcc 1500
gcctttgagg atctgagagt gagctctttt atcagaggca cccgggtggt tccacgaggt 1560
caactgtcta caagaggtgt gcagatcgcc agcaacgaga acatggagac catggatagc 1620
agcaccctgg aactgagatc cagatactgg gccatcagga cacggagcgg cggcaccacc 1680
aatcagcagc gcgccagcgc cggccagatc tctgtccagc ctacgtttag cgtgcagcgg 1740
aatttgccct tcgaacgcgc cacaatcatg gctgctttca ccggcaatac agagggcaga 1800
accagcgata tgagaacaga aattatccgt atgatggagt ccgcaaaacc tgaggacgtg 1860
tccttccaag gcagaggcgt gttcgagctg agcgacgaga aggccaccaa ccctatcgtg 1920
cctagcttcg atatgtctaa tgagggcagc tacttttttcg gagataacgc cgaagagtac 1980
gacaacatgt ctaatatgga tatcgacggc attaacaccg gcaccatcga caaaacccct 2040
gaggagatca cccctggcac cagcggcaca acccggccca tcatccgccc cgctacactg 2100
gctccaccta gcaacaagcg gaccagaaat ccctcgccag aaagagccac aacctccagc 2160
gaggacgacg tgggacggaa gacacaaaag aagcagaccc ctacagagat caagaagtct 2220
gtttacaaca tggtggtgaa actgggcgag ttctacaacc agatgatggt gaaggccggc 2280
ctgaacgacg atatggaaag aaatctgatc cagaacgccc acgccgtgga gcggattctg 2340
ctggccgcca ccgatgataa gaagaccgaa ttccagaaaa agaaaaacgc cagagacgtg 2400
aaggaaggca aggaagagat cgaccacaac aagacaggcg gcacattcta caagatggtc 2460
cgggacgaca agaccatcta cttcagccct atccggataa cattcctgaa agaagaagtg 2520
aagaccatgt acaaaaccac aatgggctct gacggcttca gcggcctgaa tcacatcatg 2580
atcggccact ctcaaatgaa cgatgtgtgc ttccagagaa gcaaggctct gaagcgcgtg 2640
ggcctggatc ctagcctgat ctctaccttc gccggcagca ccatccccag aagatcgggc 2700
gctaccggcg tggctatcaa gggaggaggc acactggtgg ctgaagccat cagattcatc 2760

ggaagagcca tggccgacag aggactcctg agagatatca aagccaaaac cgcctacgaa 2820
aaaatcctgc tgaacctgaa gaacaagtgc agcgcgcctc aacagaaggc cctggtggac 2880
caggttatcg gctctagaaa ccctggaatc gccgatatcg aggacctgac actgctggcc 2940
agatctatgg tggtggtgag accctccgtg gccagcaagg tggtgctgcc tatcagcatc 3000
tacgccaaga tccctcagct gggatttaac gtggaagaat acagcatggt tggttatgag 3060
gccatggccc tgtacaacat ggccacacct gtgtccatcc tgagaatggg cgacgatgcc 3120
aaagacaaga gccagctgtt cttcatgagc tgcttcggcg ctgcctatga ggacctgaga 3180
gtgctgtccg ctcttacagg aacagagttc aagcctagga gcgcactgaa gtgcaagggc 3240
ttccacgtgc ccgccaagga acaggtggaa ggcatgggag ctgctctgat gtccatcaag 3300
ctgcaatttt gggctcctat gacccggagc ggcggaaatg aggtgggtgg cgacggaggc 3360
agcggacaga tttcttgcag ccccgtattt gccgtggaga gaccaatcgc cctgtccaag 3420
caggccgtga gaagaatgct gagcatgaac atcgagggcc gggacgccga cgtgaagggc 3480
aacctgttga agatgatgaa cgacagcatg gccaagaaga ccagtggcaa tgccttcatc 3540
ggcaagaaga tgttccagat ctccgacaag aacaagacca accccgtgga aatccccatc 3600
aagcagacaa tccctaactt cttcttcggc agagacaccg ccgaagacta tgacgacctg 3660
gactacatag gaaattgccc aatatgggtg aaaacacctt tgaagcttgc caatggaacc 3720
aaatatagac ctcctgcaaa actattaaag gaaaggggtt tcttcggagc tattgctggt 3780
ttcctagaag gaggatggga aggaatgatt gcaggctggc acggatacac atctcacgga 3840
gcacatggag tggcagtggc ggcggacctt aagagtacgc aagaagctat aaacaagata 3900
acaaaaaatc tcaattcttt gagtgagcta gaagtaaaga atcttcaaag actaagtggt 3960
gccatggatg aactccacaa cgaaatactc gagctggatg agaaagtgga tgatctcaga 4020
gctgacacta taagctcgca aatagaactt gcagtcttgc tttccaacga aggaataata 4080
aacagtgaag atgagtga 4098

<210> 29
<211> 1365
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 29
Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1   5   10   15
Met Leu Val Ala Ser Cys Leu Gly Ile Gly Glu Cys Pro Lys Tyr Val
            20   25   30
Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile Pro Ser
        35   40   45
Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly
    50   55   60
Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn
65   70   75   80
Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala
                85   90   95
Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn
            100   105   110
Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg
        115   120   125
Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp

```
    130                 135                 140
Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu
145                 150                 155                 160
Asp Phe His Asp Ser Ala Ser Gln Gly Thr Lys Arg Ser Tyr Glu Gln
                165                 170                 175
Met Glu Thr Gly Gly Glu Arg Gln Asn Ala Thr Glu Ile Arg Ala Ser
            180                 185                 190
Val Gly Arg Met Val Gly Gly Ile Gly Arg Phe Tyr Ile Gln Met Cys
        195                 200                 205
Thr Glu Leu Lys Leu Ser Asp Tyr Glu Gly Arg Leu Ile Gln Asn Ser
    210                 215                 220
Ile Thr Ile Glu Arg Met Val Leu Ser Ala Phe Asp Glu Arg Arg Asn
225                 230                 235                 240
Lys Tyr Leu Glu Glu His Pro Ser Ala Gly Lys Asp Pro Lys Lys Thr
                245                 250                 255
Gly Gly Pro Ile Tyr Arg Arg Arg Asp Gly Lys Trp Val Arg Glu Leu
            260                 265                 270
Ile Leu Tyr Asp Lys Glu Glu Ile Arg Arg Ile Trp Arg Gln Ala Asn
        275                 280                 285
Asn Gly Glu Asp Ala Thr Ala Gly Leu Thr His Leu Met Ile Trp His
    290                 295                 300
Ser Asn Leu Asn Asp Ala Thr Tyr Gln Arg Thr Arg Ala Leu Val Arg
305                 310                 315                 320
Thr Gly Met Asp Pro Arg Met Cys Ser Leu Met Gln Gly Ser Thr Leu
                325                 330                 335
Pro Arg Arg Ser Gly Ala Ala Gly Ala Ala Val Lys Gly Val Gly Thr
            340                 345                 350
Met Val Met Glu Leu Ile Arg Met Ile Lys Arg Gly Ile Asn Asp Arg
        355                 360                 365
Asn Phe Trp Arg Gly Glu Asn Gly Arg Arg Thr Arg Ile Ala Tyr Glu
    370                 375                 380
Arg Met Cys Asn Ile Leu Lys Gly Lys Phe Gln Thr Ala Ala Gln Arg
385                 390                 395                 400
Ala Met Met Asp Gln Val Arg Glu Ser Arg Asn Pro Gly Asn Ala Glu
                405                 410                 415
Ile Glu Asp Leu Ile Phe Leu Ala Arg Ser Ala Leu Ile Leu Arg Gly
            420                 425                 430
Ser Val Ala His Lys Ser Cys Leu Pro Ala Cys Val Tyr Gly Leu Ala
        435                 440                 445
Val Ala Ser Gly Tyr Asp Phe Glu Arg Glu Gly Tyr Ser Leu Val Gly
    450                 455                 460
Ile Asp Pro Phe Arg Leu Leu Gln Asn Ser Gln Val Phe Ser Leu Ile
465                 470                 475                 480
Arg Pro Asn Glu Asn Pro Ala His Lys Ser Gln Leu Val Trp Met Ala
                485                 490                 495
Cys His Ser Ala Ala Phe Glu Asp Leu Arg Val Ser Ser Phe Ile Arg
            500                 505                 510
Gly Thr Arg Val Val Pro Arg Gly Gln Leu Ser Thr Arg Gly Val Gln
        515                 520                 525
Ile Ala Ser Asn Glu Asn Met Glu Thr Met Asp Ser Ser Thr Leu Glu
```

```
    530                 535                 540
Leu Arg Ser Arg Tyr Trp Ala Ile Arg Thr Arg Ser Gly Gly Thr Thr
545                 550                 555                 560
Asn Gln Gln Arg Ala Ser Ala Gly Gln Ile Ser Val Gln Pro Thr Phe
                565                 570                 575
Ser Val Gln Arg Asn Leu Pro Phe Glu Arg Ala Thr Ile Met Ala Ala
            580                 585                 590
Phe Thr Gly Asn Thr Glu Gly Arg Thr Ser Asp Met Arg Thr Glu Ile
        595                 600                 605
Ile Arg Met Met Glu Ser Ala Lys Pro Glu Asp Val Ser Phe Gln Gly
    610                 615                 620
Arg Gly Val Phe Glu Leu Ser Asp Glu Lys Ala Thr Asn Pro Ile Val
625                 630                 635                 640
Pro Ser Phe Asp Met Ser Asn Glu Gly Ser Tyr Phe Phe Gly Asp Asn
                645                 650                 655
Ala Glu Glu Tyr Asp Asn Met Ser Asn Met Asp Ile Asp Gly Ile Asn
            660                 665                 670
Thr Gly Thr Ile Asp Lys Thr Pro Glu Glu Ile Thr Pro Gly Thr Ser
        675                 680                 685
Gly Thr Thr Arg Pro Ile Ile Arg Pro Ala Thr Leu Ala Pro Pro Ser
    690                 695                 700
Asn Lys Arg Thr Arg Asn Pro Ser Pro Glu Arg Ala Thr Thr Ser Ser
705                 710                 715                 720
Glu Asp Asp Val Gly Arg Lys Thr Gln Lys Lys Gln Thr Pro Thr Glu
                725                 730                 735
Ile Lys Lys Ser Val Tyr Asn Met Val Val Lys Leu Gly Glu Phe Tyr
            740                 745                 750
Asn Gln Met Met Val Lys Ala Gly Leu Asn Asp Asp Met Glu Arg Asn
        755                 760                 765
Leu Ile Gln Asn Ala His Ala Val Glu Arg Ile Leu Leu Ala Ala Thr
    770                 775                 780
Asp Asp Lys Lys Thr Glu Phe Gln Lys Lys Lys Asn Ala Arg Asp Val
785                 790                 795                 800
Lys Glu Gly Lys Glu Glu Ile Asp His Asn Lys Thr Gly Gly Thr Phe
                805                 810                 815
Tyr Lys Met Val Arg Asp Asp Lys Thr Ile Tyr Phe Ser Pro Ile Arg
            820                 825                 830
Ile Thr Phe Leu Lys Glu Glu Val Lys Thr Met Tyr Lys Thr Thr Met
        835                 840                 845
Gly Ser Asp Gly Phe Ser Gly Leu Asn His Ile Met Ile Gly His Ser
    850                 855                 860
Gln Met Asn Asp Val Cys Phe Gln Arg Ser Lys Ala Leu Lys Arg Val
865                 870                 875                 880
Gly Leu Asp Pro Ser Leu Ile Ser Thr Phe Ala Gly Ser Thr Ile Pro
                885                 890                 895
Arg Arg Ser Gly Ala Thr Gly Val Ala Ile Lys Gly Gly Gly Thr Leu
            900                 905                 910
Val Ala Glu Ala Ile Arg Phe Ile Gly Arg Ala Met Ala Asp Arg Gly
        915                 920                 925
Leu Leu Arg Asp Ile Lys Ala Lys Thr Ala Tyr Glu Lys Ile Leu Leu
```

```
    930                 935                 940
Asn Leu Lys Asn Lys Cys Ser Ala Pro Gln Gln Lys Ala Leu Val Asp
945                 950                 955                 960
Gln Val Ile Gly Ser Arg Asn Pro Gly Ile Ala Asp Ile Glu Asp Leu
                965                 970                 975
Thr Leu Leu Ala Arg Ser Met Val Val Val Arg Pro Ser Val Ala Ser
            980                 985                 990
Lys Val Val Leu Pro Ile Ser Ile Tyr Ala Lys Ile Pro Gln Leu Gly
        995                 1000                1005
Phe Asn Val Glu Glu Tyr Ser Met Val Gly Tyr Glu Ala Met Ala Leu
    1010                1015                1020
Tyr Asn Met Ala Thr Pro Val Ser Ile Leu Arg Met Gly Asp Asp Ala
1025                1030                1035                1040
Lys Asp Lys Ser Gln Leu Phe Phe Met Ser Cys Phe Gly Ala Ala Tyr
                1045                1050                1055
Glu Asp Leu Arg Val Leu Ser Ala Leu Thr Gly Thr Glu Phe Lys Pro
            1060                1065                1070
Arg Ser Ala Leu Lys Cys Lys Gly Phe His Val Pro Ala Lys Glu Gln
        1075                1080                1085
Val Glu Gly Met Gly Ala Ala Leu Met Ser Ile Lys Leu Gln Phe Trp
    1090                1095                1100
Ala Pro Met Thr Arg Ser Gly Gly Asn Glu Val Gly Gly Asp Gly Gly
1105                1110                1115                1120
Ser Gly Gln Ile Ser Cys Ser Pro Val Phe Ala Val Glu Arg Pro Ile
                1125                1130                1135
Ala Leu Ser Lys Gln Ala Val Arg Arg Met Leu Ser Met Asn Ile Glu
            1140                1145                1150
Gly Arg Asp Ala Asp Val Lys Gly Asn Leu Leu Lys Met Met Asn Asp
        1155                1160                1165
Ser Met Ala Lys Lys Thr Ser Gly Asn Ala Phe Ile Gly Lys Lys Met
    1170                1175                1180
Phe Gln Ile Ser Asp Lys Asn Lys Thr Asn Pro Val Glu Ile Pro Ile
1185                1190                1195                1200
Lys Gln Thr Ile Pro Asn Phe Phe Phe Gly Arg Asp Thr Ala Glu Asp
                1205                1210                1215
Tyr Asp Asp Leu Asp Tyr Ile Gly Asn Cys Pro Ile Trp Val Lys Thr
            1220                1225                1230
Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys Leu
        1235                1240                1245
Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu Gly
    1250                1255                1260
Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly
1265                1270                1275                1280
Ala His Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala
                1285                1290                1295
Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val
            1300                1305                1310
Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu
        1315                1320                1325
Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile
```

```
    1330                1335                1340
Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile
1345                1350                1355                1360
Asn Ser Glu Asp Glu
                1365
```